STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 123594

TO: Jennifer Graser
Location: REM/3B09/3C18
Art Unit: 1645
June 7, 2004

Case Serial Number: 10/712713

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

OM nucleic - nucleic search, using sw model

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:     June  6, 2004, 04:13:01 ; Search time 135 Seconds
                                (without alignments)
                           5417.965 Million cell updates/sec

Title:          US-10-712-713-1
Perfect score:  1318
Sequence:       1 atgacacactatcatttgt........gaaaaatgcgtttaagctt 1318

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       682709 seqs, 277475446 residues

Word size :     50

Total number of hits satisfying chosen parameters:    8

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1314	99.7	2424	4	US-08-956-171E-392	Sequence 392, App
2	275	20.9	619	4	US-08-936-165A-179	Sequence 179, App
C 3	90	6.8	888	3	US-08-714-918-61	Sequence 61, Appl
C 4	90	6.8	888	3	US-09-265-315-61	Sequence 61, Appl
C 5	90	6.8	888	3	US-09-265-315-61	Sequence 61, Appl
C 6	90	6.8	888	3	US-09-266-417-61	Sequence 61, Appl
C 7	90	6.8	888	4	US-09-528-709-61	Sequence 61, Appl
C 8	90	6.8	888	4	US-09-527-745-61	Sequence 61, Appl

ALIGNMENTS

```
RESULT 1
US-08-956-171E-392
;  Sequence 392, Application US/08956171E
;  Patent No. 6593114
   GENERAL INFORMATION:
      APPLICANT: Charles Kunsch
                 Gil H. Choi
                 Patrick S. Dillon
                 Craig A. Rosen
                 Steven C. Barash
                 Michael R. Fannon
   TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
   NUMBER OF SEQUENCES: 5256
   CORRESPONDENCE ADDRESS:
```

```
            ADDRESSEE: Human Genome Sciences, Inc.
            STREET: 9410 Key West Avenue
            CITY: Rockville
            STATE: Maryland
            COUNTRY: USA
            ZIP: 20850
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
      COMPUTER: HP Vectra 486/33
      OPERATING SYSTEM: MSDOS version 6.2
      SOFTWARE: ASCII Text
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/956,171E
      FILING DATE: 20-Oct-1997
      CLASSIFICATION: <Unknown>
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 60/009,861
      FILING DATE: January 5, 1996
      APPLICATION NUMBER: 08/781,986
      FILING DATE: January 3, 1997
   ATTORNEY/AGENT INFORMATION:
      NAME: Mark J. Hyman
      REGISTRATION NUMBER: 46,789
      REFERENCE/DOCKET NUMBER: PB248P1
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (240) 314-1224
      TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 392:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 2424 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
   SEQUENCE DESCRIPTION: SEQ ID NO: 392:
US-08-956-171E-392

Query Match            99.7%;  Score 1314;  DB 4;  Length 2424;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1314;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   627 ATGACACACTATCATTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC 686

Qy    61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   687 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 746

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   747 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 806

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAAAATAGTACGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   807 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAAAATAGTACGTGCA 866

Qy   241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   867 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA 926

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   927 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 986

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   987 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGTATGGGA 1046

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1047 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 1106

Qy   481 TATAAACCTGATTACGGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTAA 540
```

NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19406-0939
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/936,165A
;   FILING DATE: 24-SEP-1997
;   CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/027,032
;   FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Gimmi, Edward R
;   REGISTRATION NUMBER: 38,891
;   REFERENCE/DOCKET NUMBER: P50549
; TELECOMUNICATION INFORMATION:
;   TELEPHONE: 610-270-4478
;   TELEFAX: 610-270-5090
;   TELEX:
; INFORMATION FOR SEQ ID NO: 179:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 619 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
US-08-936-165A-179

Query Match   20.9%;   Score 275; DB 4; Length 619;
Best Local Similarity   99.7%;   Pred. No. 1.1e-119;
Matches   325; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   611 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 552

Qy   781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 492

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 GATGTTACAAATATTAAAGAAGCATTAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 432

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 372

Qy   961 AGTGCTACAATTGAAACAGCACAGCACGCAGAAAGAAATATCCACATAAAGAGAGTTGTTGCAGTATTT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 AGTGCTACAATTGACACAGCACGACGCAGAAAGAAATATCCACATAAAGAGAGTTGTTGCAGTATT 312

Qy   1021 CAACCACACACTTTCTTAGAACACA 1046
              |||||||||||||||||||||||||
Db   311 CAACCACACACTTTCTTCTAGAACACA 286

RESULT 3
US-08-714-918-61/c
; Sequence 61, Application US/08714918
; Patent No. 6037123
; GENERAL INFORMATION:
;   APPLICANT: Benton, Bret
;   APPLICANT: Lee, Ving

Db   1107 TATAAACCTGATTACGCAATTAATGACAAATATGATTTCGATCATCCTGATTATTTTAAA 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAGAAATGGCACATAATGTTAAAAAGGTATT 600

Db   1167 GATATTAATGATGTTTTGATGCATTCCAAGAGAAATGGCACATAATGTTAAAAAGGTATT 1226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 ATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGATGTTCAATTTATTAT 660

Db   1227 ATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGATGTTCCAATTTATTAT 1286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 TATGGATTAAAGATTCGGATGACAATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720

Db   1287 TATGGATTAAAGATTCGGATGACAATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 1346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780

Db   1347 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 1406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 840

Db   1407 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 1466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 900

Db   1467 GATGTTACAAATATTAAAGAAGCATTAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTTCAAT 1526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Db   1527 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 1586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 AGTGCTACAATTGAAACAGCACAGCACGCAGAAAGAAATATCCACATAAAGAGAGTTGTTGCAGTATTT 1020

Db   1587 AGTGCTACAATTGAAACAGCACAGCACGCAGAAAGAAATATCCACATAAAGAGAGTTGTTGCAGTATT 1646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 CAACCACACACTTTCTTAGAACACAGGCATTTTTAAATGAATTGCAGAAAGTTTAAGT 1080

Db   1647 CAACCACACACTTTCTTCTAGAACACAGGCATTTTTAAATGAATTGCAGAAAGTTTAAGT 1706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGAAATTTTTGGATCAATTAGAGAAAATACTCGCGCA 1140

Db   1707 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGAAATTTTTGGATCAATTAGAGAAAATACTCGCGCA 1766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200

Db   1767 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 ATTAATGTATTAGAACAATTGATAATGCTGTATTTATTTATTATTGGGTGCAGGTGATATT 1260

Db   1827 ATTAATGTATTAGAACAATTGATAATGCTGTATTTATTTATTATTGGGTGCAGGTGATATT 1886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGCCATGAAAAATGCGTTTTAA 1314

Db   1887 CAAAAATTACAAAATGCATATTTAGATAAATTAGCCATGAAAAATGCGTTTTAA 1940

RESULT 2
US-08-936-165A-179/c
; Sequence 179, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
;   APPLICANT: Black, Michael
;   APPLICANT: Burnham, Martin
;   APPLICANT: Hodgson, John
;   APPLICANT: Knowles, David
;   APPLICANT: Lonetto, Michael
;   APPLICANT: Nicholas, Richard
;   APPLICANT: Pratt, Julie
;   APPLICANT: Reichard, Richard
;   APPLICANT: Rosenberg, Martin
;   APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582e1 Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses

~ APPLICANT: Malouin, Francois
~ APPLICANT: Martin, Patrick K.
~ APPLICANT: Schmid, Molly B.
~ APPLICANT: Sun, Dongxu
~ TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
~ TITLE OF INVENTION: TARGET GENES
~ NUMBER OF SEQUENCES: 111
~ CORRESPONDENCE ADDRESS:
~ ADDRESSEE: Lyon & Lyon
~ STREET: 633 West Fifth Street
~ STREET: Suite 4700
~ CITY: Los Angeles
~ STATE: California
~ COUNTRY: U.S.A.
~ ZIP: 90071-2066
~ COMPUTER READABLE FORM:
~ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
~ MEDIUM TYPE: storage
~ COMPUTER: IBM Compatible
~ OPERATING SYSTEM: IBM P.C. DOS 5.0
~ SOFTWARE: Word Perfect 5.1
~ CURRENT APPLICATION DATA:
~ APPLICATION NUMBER: US/08/714,918
~ FILING DATE: September 13, 1996
~ CLASSIFICATION: 424
~ PRIOR APPLICATION DATA:
~ APPLICATION NUMBER: 60/009,102
~ FILING DATE: December 22, 1995
~ APPLICATION NUMBER: 60/003,798
~ FILING DATE: September 15, 1995
~ ATTORNEY/AGENT INFORMATION:
~ NAME: Warburg, Richard J.
~ REGISTRATION NUMBER: 32,327
~ REFERENCE/DOCKET NUMBER: 222/005
~ TELECOMMUNICATION INFORMATION:
~ TELEPHONE: (213) 489-1600
~ TELEFAX: (213) 955-0440
~ TELEX: 67-3510
~ INFORMATION FOR SEQ ID NO: 61:
~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 888 base pairs
~ TYPE: nucleic acid
~ STRANDEDNESS: single
~ TOPOLOGY: linear
US-08-714-918-61

Query Match   6.8%;   Score 90; DB 3; Length 888;
Best Local Similarity 100.0%;   Pred. No. 8.6e-33;
Matches   90; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1149 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCTATTAATGT 1208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCTATTAATGT 508

Qy   1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
          |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478

RESULT 4
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
; GENERAL INFORMATION:
~ APPLICANT: Benton, Bret
~ APPLICANT: Lee, Ving J.
~ APPLICANT: Malouin, Francois
~ APPLICANT: Martin, Patrick K.
~ APPLICANT: Schmid, Molly B.
~ APPLICANT: Sun, Dongxu
~ TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
~ TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
~ TITLE OF INVENTION: TARGET GENES

~ NUMBER OF SEQUENCES: 111
~ CORRESPONDENCE ADDRESS:
~ ADDRESSEE: Lyon & Lyon
~ STREET: 633 West Fifth Street
~ STREET: Suite 4700
~ CITY: Los Angeles
~ STATE: California
~ COUNTRY: U.S.A.
~ ZIP: 90071-2066
~ COMPUTER READABLE FORM:
~ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
~ MEDIUM TYPE: storage
~ COMPUTER: IBM Compatible
~ OPERATING SYSTEM: IBM P.C. DOS 5.0
~ SOFTWARE: Word Perfect 5.1
~ CURRENT APPLICATION DATA:
~ APPLICATION NUMBER: US/09/265,315
~ FILING DATE: March 9, 1999
~ CLASSIFICATION: 435
~ PRIOR APPLICATION DATA:
~ APPLICATION NUMBER: 08/714,918
~ FILING DATE: September 13, 1996
~ APPLICATION NUMBER: 60/009,102
~ FILING DATE: December 22, 1995
~ APPLICATION NUMBER: 60/003,798
~ FILING DATE: September 15, 1995
~ ATTORNEY/AGENT INFORMATION:
~ NAME: Warburg, Richard J.
~ REGISTRATION NUMBER: 32,327
~ REFERENCE/DOCKET NUMBER: 240/247
~ TELECOMMUNICATION INFORMATION:
~ TELEPHONE: (213) 489-1600
~ TELEFAX: (213) 955-0440
~ TELEX: 67-3510
~ INFORMATION FOR SEQ ID NO: 61:
~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 888 base pairs
~ TYPE: nucleic acid
~ STRANDEDNESS: single
~ TOPOLOGY: linear
US-09-265-315-61

Query Match   6.8%;   Score 90; DB 3; Length 888;
Best Local Similarity 100.0%;   Pred. No. 8.6e-33;
Matches   90; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1149 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCTATTAATGT 1208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCTATTAATGT 508

Qy   1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
          |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478

RESULT 5
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
; GENERAL INFORMATION:
~ APPLICANT: Benton, Bret
~ APPLICANT: Lee, Ving J.
~ APPLICANT: Malouin, Francois
~ APPLICANT: Martin, Patrick K.
~ APPLICANT: Schmid, Molly B.
~ APPLICANT: Sun, Dongxu
~ TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
~ TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
~ TITLE OF INVENTION: TARGET GENES
~ NUMBER OF SEQUENCES: 111
~ CORRESPONDENCE ADDRESS:
~ ADDRESSEE: Lyon & Lyon
~ STREET: 633 West Fifth Street

```
            STREET: Suite 4700
            CITY: Los Angeles
            STATE: California
            COUNTRY: U.S.A.
            ZIP: 90071-2066
       COMPUTER READABLE FORM:
            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
            MEDIUM TYPE: storage
            COMPUTER: IBM Compatible
            OPERATING SYSTEM: IBM P.C. DOS 5.0
            SOFTWARE: Word Perfect 5.1
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/09/265,315
            FILING DATE: March 9, 1999
            CLASSIFICATION: 435
       PRIOR APPLICATION DATA:
            APPLICATION NUMBER: 08/714,918
            FILING DATE: September 13, 1996
            APPLICATION NUMBER: 60/009,102
            FILING DATE: December 22, 1995
            APPLICATION NUMBER: 60/003,798
            FILING DATE: September 15, 1995
       ATTORNEY/AGENT INFORMATION:
            NAME: Warburg, Richard J.
            REGISTRATION NUMBER: 32,327
            REFERENCE/DOCKET NUMBER: 240/247
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (213) 489-1600
            TELEFAX: (213) 955-0440
            TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
       SEQUENCE CHARACTERISTICS:
            LENGTH: 888 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
US-09-265-315-61
```

```
Query Match             6.8%;   Score 90;  DB 3;  Length 888;
Best Local Similarity 100.0%;   Pred. No. 8.6e-33;
Matches   90;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  1149 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCTATTAATGT 1208
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCTATTAATGT 508

Qy  1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
         |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478
```

```
RESULT 6
US-09-266-417-61/c
; Sequence 61, Application US/09266417
; Patent No. 6228588
; GENERAL INFORMATION:
       APPLICANT: Benton, Bret
       APPLICANT: Lee, Ving J.
       APPLICANT: Malouin, Francois
       APPLICANT: Martin, Patrick K.
       APPLICANT: Schmid, Molly B.
       APPLICANT: Sun, Dongxu
       TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
       TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
       TITLE OF INVENTION: TARGET GENES
       NUMBER OF SEQUENCES: 111
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Lyon & Lyon
            STREET: 633 West Fifth Street
            STREET: Suite 4700
            CITY: Los Angeles
            STATE: California
            COUNTRY: U.S.A.
            ZIP: 90071-2066
       COMPUTER READABLE FORM:
            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
            MEDIUM TYPE: storage
            COMPUTER: IBM Compatible
            OPERATING SYSTEM: IBM P.C. DOS 5.0
            SOFTWARE: Word Perfect 5.1
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/09/266,417
            FILING DATE: March 9, 1999
            CLASSIFICATION: 435
       PRIOR APPLICATION DATA:
            APPLICATION NUMBER: 08/714,918
            FILING DATE: September 13, 1996
            APPLICATION NUMBER: 60/009,102
            FILING DATE: December 22, 1995
            APPLICATION NUMBER: 60/003,798
            FILING DATE: September 15, 1995
       ATTORNEY/AGENT INFORMATION:
            NAME: Warburg, Richard J.
            REGISTRATION NUMBER: 32,327
            REFERENCE/DOCKET NUMBER: 240/248
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (213) 489-1600
            TELEFAX: (213) 955-0440
            TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
       SEQUENCE CHARACTERISTICS:
            LENGTH: 888 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
US-09-266-417-61
```

```
Query Match             6.8%;   Score 90;  DB 3;  Length 888;
Best Local Similarity 100.0%;   Pred. No. 8.6e-33;
Matches   90;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  1149 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCTATTAATGT 1208
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCTATTAATGT 508

Qy  1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
         |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478
```

```
RESULT 7
US-09-528-709-61/c
; Sequence 61, Application US/09528709
; Patent No. 6630303
; GENERAL INFORMATION:
       APPLICANT: Benton, Bret
                  Lee, Ving
                  Malouin, Francois
                  Martin, Patrick K.
                  Schmid, Molly B.
                  Sun, Dongxu
       TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
                           TARGET GENES
       NUMBER OF SEQUENCES: 111
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Lyon & Lyon
            STREET: 633 West Fifth Street
                    Suite 4700
            CITY: Los Angeles
            STATE: California
            COUNTRY: U.S.A.
            ZIP: 90071-2066
       COMPUTER READABLE FORM:
            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
                         storage
            COMPUTER: IBM Compatible
```

```
              OPERATING SYSTEM: IBM P.C. DOS 5.0
              SOFTWARE: Word Perfect 5.1
         CURRENT APPLICATION DATA:
              APPLICATION NUMBER: US/09/528,709
              FILING DATE: 17-Mar-2000
              CLASSIFICATION: 424
         PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US/08/714,918
              FILING DATE: September 13, 1996
              APPLICATION NUMBER: 60/009,102
              FILING DATE: December 22, 1995
              APPLICATION NUMBER: 60/003,798
              FILING DATE: September 15, 1995
         ATTORNEY/AGENT INFORMATION:
              NAME: Warburg, Richard J.
              REGISTRATION NUMBER: 32,327
              REFERENCE/DOCKET NUMBER: 222/005
         TELECOMMUNICATION INFORMATION:
              TELEPHONE: (213) 489-1600
              TELEFAX: (213) 955-0440
              TELEX: 67-3510
    INFORMATION FOR SEQ ID NO: 61:
         SEQUENCE CHARACTERISTICS:
              LENGTH: 888 base pairs
              TYPE: nucleic acid
              STRANDEDNESS: single
              TOPOLOGY: linear
         SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-528-709-61

Query Match          6.8%;  Score 90;  DB 4;  Length 888;
Best Local Similarity  100.0%;  Pred. No. 8.6e-33;
Matches   90;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  1149 ACAAGATTTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGAATTCTATTAATGT 1208
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGAATTCTATTAATGT 508

Qy  1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
         |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478


RESULT 8
US-09-527-745-61/c
; Sequence 61, Application US/09527745
; Patent No. 6638718
GENERAL INFORMATION:
     APPLICANT: Benton, Bret
              Lee, Ving
              Malouin, Francois
              Martin, Patrick K.
              Schmid, Molly B.
              Sun, Dongxu
     TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
                         TARGET GENES
     NUMBER OF SEQUENCES: 111
     CORRESPONDENCE ADDRESS:
         ADDRESSEE: Lyon & Lyon
         STREET: 633 West Fifth Street
                 Suite 4700
         CITY: Los Angeles
         STATE: California
         COUNTRY: U.S.A.
         ZIP: 90071-2066
     COMPUTER READABLE FORM:
         MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
                      storage
         COMPUTER: IBM Compatible
         OPERATING SYSTEM: IBM P.C. DOS 5.0
         SOFTWARE: Word Perfect 5.1
     CURRENT APPLICATION DATA:
         APPLICATION NUMBER: US/09/527,745
```

```
              FILING DATE: 17-Mar-2000
              CLASSIFICATION: 424
         PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US/08/714,918
              FILING DATE: September 13, 1996
              APPLICATION NUMBER: 60/009,102
              FILING DATE: December 22, 1995
              APPLICATION NUMBER: 60/003,798
              FILING DATE: September 15, 1995
         ATTORNEY/AGENT INFORMATION:
              NAME: Warburg, Richard J.
              REGISTRATION NUMBER: 32,327
              REFERENCE/DOCKET NUMBER: 222/005
         TELECOMMUNICATION INFORMATION:
              TELEPHONE: (213) 489-1600
              TELEFAX: (213) 955-0440
              TELEX: 67-3510
    INFORMATION FOR SEQ ID NO: 61:
         SEQUENCE CHARACTERISTICS:
              LENGTH: 888 base pairs
              TYPE: nucleic acid
              STRANDEDNESS: single
              TOPOLOGY: linear
         SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-527-745-61

Query Match          6.8%;  Score 90;  DB 4;  Length 888;
Best Local Similarity  100.0%;  Pred. No. 8.6e-33;
Matches   90;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  1149 ACAAGATTTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGAATTCTATTAATGT 1208
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 ACAAGATTTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGAATTCTATTAATGT 508

Qy  1209 ATTAGAACAATTTGATAATGCTGTTATTTT 1238
         |||||||||||||||||||||||||||||
Db   507 ATTAGAACAATTTGATAATGCTGTTATTTT 478

Search completed: June  6, 2004, 07:10:16
Job time : 136 secs
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:      June  6, 2004, 04:34:41 ; Search time 315 Seconds
                                        (without alignments)
                                        5591.218 Million cell updates/sec

Title:       US-10-712-713-1
Perfect score: 1318
Sequence:    1 atgacacactcatttgt.........gaaaaatgcgttttaagctt 1318

Scoring table: OLIGO_NUC
             Gapop 60.0 , Gapext 60.0

Searched:    2089662 seqs, 668146292 residues

Word size :  50

Total number of hits satisfying chosen parameters:       0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Pending Patents_NA_New:*
             1: /cgn2_6/ptodata72/pna/PCT_NEW_COMB.seq:*
             2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
             3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
             4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
             5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
             6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
             7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES

                  %
Result       Query
No.   Score  Match Length DB  ID                              Description
------------------------------------------------------------------------------

No matches found

Search completed: June  6, 2004, 08:36:08
Job time : 315 secs
```

Search completed: June   6 , 2004, 07:07:51
Job time : 3824 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
        GenCore version 5.1.6

Run on: June   6, 2004, 04:07:36 ; Search time 3824 Seconds
                                         (without alignments)
                                         10292.457 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacactatcatttgt.........gaaaaatgcgtttaagctt 1318

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 1491090276 residues

Word size :   50

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   EST:*
                1: em_estba:*
                2: em_esthum:*
                3: em_estin:*
                4: em_estmu:*
                5: em_estov:*
                6: em_estpl:*
                7: em_estro:*
                8: em_htc:*
                9: gb_est1:*
               10: gb_est2:*
               11: gb_htc:*
               12: gb_est3:*
               13: gb_est4:*
               14: gb_est5:*
               15: em_estfun:*
               16: em_eston:*
               17: em_gss_hum:*
               18: em_gss_inv:*
               19: em_gss_pln:*
               20: em_gss_vrt:*
               21: em_gss_fun:*
               22: em_gss_mam:*
               23: em_gss_mus:*
               24: em_gss_pro:*
               25: em_gss_rod:*
               26: em_gss_phg:*
               27: em_gss_vrl:*
               28: gb_gss1:*
               29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                 SUMMARIES

Result   Query
No.   Score Match Length DB ID   Description

No matches found

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    June  5, 2004, 23:40:14 ; Search time 3824 Seconds
                                    (without alignments)
                          10292.457 Million cell updates/sec

Title:         US-10-712-713-1
Perfect score: 1318
Sequence:      1 atgacacactatcatttgt.........gaaaatgcgttttaagctt 1318

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      27513289 seqs, 1493109027б residues

Total number of hits satisfying chosen parameters:     55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
            1:  em_estba:*
            2:  em_esthum:*
            3:  em_estin:*
            4:  em_estmu:*
            5:  em_estov:*
            6:  em_estpl:*
            7:  em_estro:*
            8:  em_htc:*
            9:  gb_est1:*
            10: gb_est2:*
            11: gb_htc:*
            12: gb_est3:*
            13: gb_est4:*
            14: gb_est5:*
            15: em_estfun:*
            16: em_estom:*
            17: em_gss_hum:*
            18: em_gss_inv:*
            19: em_gss_pln:*
            20: em_gss_vrt:*
            21: em_gss_fun:*
            22: em_gss_mam:*
            23: em_gss_mus:*
            24: em_gss_pro:*
            25: em_gss_rod:*
            26: em_gss_phg:*
            27: em_gss_vrl:*
            28: gb_gss1:*
            29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description	
c 1	373.6	28.3	2937	28	BH771016	BH771016 LlMGtag73
c 2	90	6.8	728	28	BH384961	BH384961 AG-ND-161
c 3	85.8	6.5	584	28	BH400896	BH400896 AG-ND-145
c 4	81.8	6.2	1101	29	CNS0039G	AL063921 Drosophil
5	77.8	5.9	1200	13	BX415878	BX415878
6	77.8	5.9	1200	13	BX437758	BX437758
7	75.4	5.7	1101	29	CNS00EVL	AL069706 Drosophil
8	75	5.7	908	28	AZ548467	ENTEK30TR
9	71.8	5.4	1099	13	BX456575	BX456575
c 10	70.4	5.3	849	28	AZ546009	ENTFW53TF
c 11	69.6	5.3	754	12	BJ376728	BJ376728
c 12	68.6	5.2	1056	13	BX415058	BX415058
c 13	68.2	5.2	594	13	BU498120	PfESToab9
14	67.8	5.1	1201	9	ALS36104	ALS36104
15	67.2	5.1	844	29	BX139987	Danio rer
c 16	65.8	5.0	905	28	AZ550256	ENTEV58TR
c 17	65.6	5.0	1200	13	BX415878	BX415878
18	65.2	4.9	552	9	AU269872	AU269872
c 19	64.6	4.9	825	12	BJ374097	BJ374097
20	64.6	4.9	1201	13	BX439779	BX439779
21	64.4	4.9	806	29	CNS04AEE	AL281759 Tetraodon
22	64	4.9	479	28	BH211305	Sm1-50G22
c 23	63.8	4.8	843	28	AZ551618	ENTDV54TR
24	63.6	4.8	751	12	BJ374223	BJ374223
c 25	63.4	4.8	1200	13	BX437758	BX437758
c 26	63.2	4.8	608	12	BJ330059	BJ330059
c 27	62.6	4.7	856	28	BH178274	011_F_07-
c 28	62.6	4.7	856	29	CNS07KML	T7_end_of
c 29	62.2	4.7	987	29	CNS014PQ	AL104456 Drosophil
c 30	61.8	4.7	829	29	BX173672	Danio rer
c 31	61.8	4.7	1061	13	BX437039	BX437039
c 32	61.6	4.7	900	28	AZ549980	ENTDD94TF
33	61.4	4.7	1098	13	BX377526	BX377526
c 34	61.4	4.7	1201	13	BX366070	BX366070
c 35	61.2	4.6	541	29	CNS04KUS	AL295309 Tetraodon
c 36	61.2	4.6	1391	29	CG754863	CG754863 P050-2-G0
c 37	61	4.6	1201	13	BX406288	BX406288
c 38	61	4.6	1201	13	BX461824	BX461824
c 39	60.8	4.6	548	28	BH200391	Sm1-39D16
c 40	60.8	4.6	1201	13	BX355654	BX355654
41	60.4	4.6	313	28	BH204978	Sm1-4319.
42	60.4	4.6	795	28	AZ528485	ENTCM64TF
c 43	60.4	4.6	823	28	AZ676218	ENTKE36TR
44	60.4	4.6	843	28	BH139532	ENTNG88TF
45	60.4	4.6	996	29	CNS00FUH	AL071063 Drosophil

ALIGNMENTS

```
RESULT 1
BH771016/c
LOCUS       BH771016               2937 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LlMGtag738 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, genomic survey sequence.
ACCESSION   BH771016
VERSION     BH771016.1  GI:20373973
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 2937)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is murC (91%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 2907.
```

FEATURES   Location/Qualifiers
     source   1..2937
                    /organism="Lactococcus lactis subsp. cremoris"
                    /mol_type="genomic DNA"
                    /strain="MG1363"
                    /sub_species="cremoris"
                    /db_xref="taxon:1359"
                    /clone_lib="MG1363 Random Sequence Tag Library"
                    /note="Vector: pSGMU2; Site 1: SmaI; Library of
                    chromosomal fragments of L.lactis strain MG1363 was
                    prepared by partial AluI digestion or by sonication."
ORIGIN

Query Match   28.3%;   Score 373.6;   DB 28;   Length 2937;
Best Local Similarity   59.4%;   Pred. No. 4e-68;
Matches   671;   Conservative   0;   Mismatches 454;   Indels   5;   Gaps   2;

Qy   10 TATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATCATGCATGAT   69
             ||||||||| |||||||||||| ||||| |||||||||| |||||||| ||||||||||
Db   1174 TATCATTTACCCGTATTAAAGGTTCAGGCATGAGTGCACTTGCATTGATTGTACATCAA 1115

Qy   70 TTAGGACATGAAGTTCAAGGATCCGATATTGAGAACTACGTATTTACAGAAGTTGCTCTT 129
             ||||||| ||||||||||||| | ||||| ||||||||| ||||||||||||||||||
Db   1114 ATGGGCAAAAACGTTCAAGGTTCAGATTCAACAGACTACTTTTTTACCAACGTGGTCTT 1055

Qy   130 AGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAAGATATGGTA 189
             || ||||||||| ||||||||| |||||||| ||||||||||| || |||||||||||
Db   1054 GAACAAGCGGGTGTTCCACTTTTTGCCGTTTGACGGAAAAAAATATTTAAACCTGAGTTTGAA 995

Qy   190 GTTATACAAGGTAATGCATT--CGGGAGTAGCCATGAGAGAAATAGTACGTGCACATCAA 246
             || ||||| |||||||||| |||||||||||||| |||||||| |||||| ||||||
Db   994 TTGATTGTGGGAATGCTTTTCGTGATGATAATAATGTCGAAATTGCTTTCGCGCCATAAA 935

Qy   247 TTGAAATTAGATGTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAATATACT 306
             |||||||||| || |||||| | |||| |||||||||| |||| ||| |||||||
Db   934 AATGGCTTCCCATTCCAACGTTATCACGTATCATGAATTTTTGGGCCATTCATGTAAGATTCCCA 875

Qy   307 TCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGG--TTTATTATCACATG 364
             |||||||| |||||||||||||||||||||||||||||||||| ||||| |||||||
Db   874 AGTATTGGGGTTGCAGGCGCACTCGGTAAGACTTCAACCAATGGAAAATGTTAGGTTCATG 815

Qy   365 TTAATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGCTATGGGATTGC 424
             ||||||||||||||| |||||||||||||| ||||||||| ||| ||||||||| ||
Db   814 TGAAGTCGAAAATCGGGGACATTTCTTACCTAATTGGTTATGGAACTCGGGCGTGGAAATG 755

Qy   425 CTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGTATA 484
             |||||||| ||||| ||| ||||||||||| |||||||||||||| |||||| |||||
Db   754 TCGAAAGTGAATATTTTTGCTTTGAATCTAATCTAATGAGTAGTATGAACGTCATTCATGCCCATATC 695

Qy   485 AACCTGATTACGCAATTATGACAAATAATTGATTTCGATCATCCTGATTATTTTAAAGATA 544
             ||||||||||| ||||| |||||||||||| |||||||| |||||||| |||| |||||
Db   694 ATCTTGAATATACCATTATGACAAATATTGACTTTGACCACCCTGACTATTTTGAAGGAA 635

Qy   545 TTAATGAGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATTATTG 604
             ||||| ||||||| ||||||||||||||||||||| ||||| |||||||||||||||||
Db   604 TTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGAATTCTTTG 575

Qy   605 CTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTATTATG 664
             |||||||||||| |||||||| ||| ||||||| |||| |||||||| ||||| |||||
Db   574 CTTATGGTGAAGATGTGAATTGCTCGCTAAAATTGTCTGTCTAAAGCGCCTATTTATTATTATG 515

Qy   665 GATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGTACTG 724
             ||||| ||| ||| |||| ||||||||||| |||| |||| ||||||||||||||||||
Db   514 GTTTTGAAGTAATGATGATTACCGTGCAGTGAGTTTTATGATCACTTCCGTCTCCACAATATGGTG 455

Qy   725 CTTTTGATGTATGTGGATGGTGAGTTTATGATCACTTCCGTCTCCACAATATGGTG 784
             ||||||||||| ||||||| ||||| |||||||||| ||||||||| |||||
Db   454 CTTTTGATGCCTATTTCCGTGGTGAAAAAATTGGTCATTTTGTTGTTCCTGCCATGGAA 395

Qy   785 ACCATACAGTTTAAATGCCATTAGTCGTGTAATTGCCGATTAGTTATTTAGAGAAGCTAGATG 844
             || |||| ||||||||||| ||||| |||| |||||| ||| ||||||||||||||| ||
Db   394 AACATAAATGTCTTGAATGCTTTATCTGTTGTGGGCCGTTGTGGGCCGATAATCTTGGTTTCGGATA 335

to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

```
FEATURES             Location/Qualifiers
     source          1..584
                     /organism="Anopheles gambiae"
                     /mol_type="genomic DNA"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-145A13"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
```

```
ORIGIN

Query Match               6.5%;   Score 85.8;  DB 28;  Length 584;
Best Local Similarity   52.4%;   Pred. No. 8.6e-08;
Matches 189;  Conservative  0;   Mismatches 172;  Indels  0;   Gaps  0;

Qy   920 AAGTTATTGTAGATGATTATGCACACATCCAAGAGAAATTAGTGTCACAATTGAAACAG  979
         ||||| ||||||||||||| ||| |||| ||  | || || ||| ||||||||||| |||
Db   524 AAATATATATAGATGACTATGCCCATCACCCAACTGAGCTAAATGCTGTAATTGGTTCCA  465

Qy   980 CACGAAAGAAATATCCACATAAAGAGAGTTGTTGCAGTATTTCAACCACACACTTTCTCTA 1039
         ||  ||  | | ||| |  | |||||  ||| | |  | ||||||||||   |  |  |||
Db   464 TAAAAACCTTTTATCCCGATAAAAAGTCGTTAATAGTCTTTCAACCGCATTTATTTACCA   405

Qy  1040 GAACACAGGCATTTTAAATGAATTGCAGAAAGTTAAGTAAAGCAGATCGTGTATTCT     1099
         ||||| | | | |||  | | ||| | |||||| |||||||| || | |||| ||||
Db   404 GAACAAGAGACGATTTTGTAGATGGTTTTGCTGAAAGTTGAGTAAAGGAGAAGAGTTGATTT  345

Qy  1100 TATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCATTAACGATACAAAGATTTAA 1159
         ||| |||| | | |||| |  ||||||| || | | || |  || |   || | ||  |||
Db   344 TATTAGATATATATCCAGCTCGTGAATTGCCAATAGAAGGTGTGACTTCTCGATTGGTTAT   285

Qy  1160 TTGATAAAATTGAAGGTGCATCGTTAATTAATTAATGAAGATTCCATTAATGTGTATTAGAACAAT 1219
         | | | || ||||| | ||  |    || |||| | ||| | |  |  ||||| ||| ||||||
Db   284 TGGAAAAGTAAAAAGTGAAAGAAGAATAAAGCAGCCTTCAGGATGCATTTGAAAAA       225

Qy  1220 TTGATAATGCTGTTATTTTTATTATTATTATGGGTGCAGTGATATTCAAAAATTACAAAATGCAT 1279
         || | || | | |   || |  | |||| |  | |  || |||| | |||||| ||| ||||||
Db   224 TTAAAATAAAGAATTTGATATTCTCCTTACAGTCGGAGCAGAAATATAGATACTTTGT        165

Qy  1280 A 1280
         |
Db   164 A 164
```

RESULT 4

CNS0039G/c
```
LOCUS       CNS0039G               1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
```

```
Query Match               6.8%;   Score 90;  DB 28;  Length 728;
Best Local Similarity   51.6%;   Pred. No. 1.1e-08;
Matches 232;  Conservative  0;   Mismatches 215;  Indels  3;   Gaps  1;

Qy   834 GAAGCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCG   893
         | ||| ||| |||  || ||   |||||| ||||| ||  || | || ||| |||| ||
Db   624 GGAGCAAGCTATGAGAAGATCTTAAAAAAGGAATTAAATCTTTCAAAGGGATTAAAAGAAG  565

Qy   894 TTTCAATGAAACTACAATTGCAAAT--CAAGTTATTGTAGATGATTATGCACACCATCC   950
         | | ||  |||||  | ||| |||    |||| ||||||||||| ||||| |  || |||
Db   564 ATATACCAGCATTTCTTTGAGAATGGGAAAATATATAGATGACTATGCCCATCACCCC    505

Qy   951 AAGAGAAATTAGTGTCACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAGAGTTGT 1010
         || |||| ||||| ||| | || ||  |  |  ||  | | ||| |  | |||| | |||
Db   504 AACTGAGCTAAATGCTGTAATTGGTTCCATAAAAACCTTTTATCCCGATAAAAAGCTGTT   445

Qy  1011 TGCAGTATTTCAACCACACACTTTCTCTAGAACACAGGCATTTTTAAATGAATTGCAGA   1070
         | | | |||||||||| |  | |  |||||||| | | | ||||| | | ||| | |||
Db   444 AATAGTCTTTCAACCGCATTTATTACCAGAACAAGAGACGATTTTGTAGATGTTTTGCTGA  385

Qy  1071 AAGTTTAAGTAAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAA  1130
         ||||| |||||||| || | |||| || | | |||| | | |||| |  |||||||| ||
Db   384 AAGTTTGAGTAAAGGAGAAGAGTTGATTTTATTAGATATATATCCAGCTCGTGAATTGCC   325

Qy  1131 TACTGGCGCATTAACGATACAAGATTTAAATTGATAAAATTGAAGGTGCATCGTTAATTAA 1190
         |  | | |  || | | || | || |||| | | | || ||||| | ||  |   || |||
Db   324 AATAGAAGGTGTGACTTCTCGATTGGTTATTGGAAAAGTAAAAAGTGAAAGAAGAAGAAAT  265

Qy  1191 TGAAGATTCCATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTTATGGGTGC   1250
         || | | |  |  ||||  || |||||| || | || | | |   || ||| | |  |
Db   264 AAGCAGCCTTCAGGATGCATTTGAAAAATTAAAAATAAAGAATTTGATATTCTCCTTAC    205

Qy  1251 AGGTGATATTCAAAAATTACAAAATGCATA 1280
         || | | |  || |||| ||| ||| || |
Db   204 AGTCGGAGCAGGAAATATAGATACTTTGTA  175
```

RESULT 3

BH400896/c
```
LOCUS       BH400896                584 bp    DNA     linear   GSS 11-DEC-2001
DEFINITION  AG-ND-145A13.TF ND-TAM Anopheles gambiae genomic clone
            AG-ND-145A13, genomic survey sequence.
ACCESSION   BH400896
VERSION     BH400896.1  GI:17347112
KEYWORDS    GSS.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
            Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
            Gardner,M.J. and Collins,F.H.
  TITLE     Construction of a BAC library and generation of BAC end
            sequence-tagged connectors for genome sequencing of the African
            malaria mosquito Anopheles gambiae
  JOURNAL   Mol. Genet. Genomics 268 (6), 720-728 (2003)
   MEDLINE  22542063
    PUBMED  12655398
COMMENT     Other GSSs: AG-ND-145A13.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
```

Db   440 TTTWTWTTWWAATTATTTWTTWTTWTAAAAAAAAWWWWTWTATAT 396

RESULT 5
BX415878
LOCUS   BX415878   1200 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION   BX415878 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YI04
             5-PRIME, mRNA sequence.
ACCESSION   BX415878
VERSION   BX415878.1 GI:30765550
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1200)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0CAP008BB02QP1.
FEATURES   Location/Qualifiers
     source 1..1200
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0CAP008YI04"
             /tissue_type="THYMUS"
             /clone_lib="Homo sapiens THYMUS"
             /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
             with a NotI-oligo(dT) primer. Five prime end enriched,
             double-strand cDNA was digested with Not I and cloned into
             the Not I and EcoRV sites of the pCMVSPORT 6 vector.
             Library was not normalized."
ORIGIN
Query Match   5.9%; Score 77.8; DB 13; Length 1200;
Best Local Similarity 33.9%; Pred. No. 3.9e-06;
Matches 282; Conservative 115; Mismatches 430; Indels 4; Gaps 1;

Qy 417 GGGATTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTT 476
        ||||| :|||||||| :||| ||| ||| ||:: : ||::: |||||: ||||||||
Db 369 GGGGKKKGGGGGGGGGDWWWWTTTWTTWHMAAADCTTAWTWTTWHCCATSTTTSCTTTT 428

Qy 477 AAGTTATAAACCTGATTACGCAATTATGACAAATATGATTCGATCATCCTGATTATTT 536
        ||:::|| : |||||||||: ||||| ||||:|||:||| ||||:||| |||||||
Db 429 AMAAATTTBBSSTWATATTTTATTTTTATTAATTTATTWAWTTTATTATTTTTTAA 488

Qy 537 TAAAGATATTAATAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGG 596
        ||:|:|||||||:||||: ||| |||| ::||| ||||::| | || :|||:||| |||||
Db 489 AAATTWTTTAATTTTAAAAWTTWAAWWTTWAWATTTTTAATTWAAATTAATWATT 548

Qy 597 TATTAATTGCTTGGGGTGATGATGAACATCTACGTAAAATTCAGATGTTCCAATTTA 656
        |||| | :|||::||||| | ||:| || :|:| |||| ||:| || || ||||
Db 549 TTTTTTTAAATATTTTAAAAWAAATWAATTWWTTTTWTWAAAWWTTTTATWTTTTTW 608

Qy 657 TTATTATTGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAA 716
        | |:| ||||||||| ||||| | |:||||||:||||| | ||||| | |||| | |:
Db 609 TWTTWTTTATTTTTTATTAATWTTTTAATAATWAATATAAATATTAAWTWRAWWAAWT 668

Qy 717 AGGTACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATTATTATTTTAATTWWTTTTAA 776
        ||| | | ||||||||:|:| :||| | :| |||||||||| ||| ||||| |||||
Db 669 WATATWTTTTTTTTAATTAGTTAATTATATATTATTTTATTWTTTAATWWWWATTTTT 728

Qy 777 ATATGGTGACCATACAGTTTAAATGCATTAGCTGTATTGCCATTAGTTATTTAGAGAA 836
        :|||||| || |||| :||||||||||||||:||||| : :|||||||||||||||
Db 729 WAYAAGAAAWTATTAATACATNATATATATATAWDCATAAAATATWAWTWAWTAAAATA 788

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp; the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source 1..1101
             /organism="Drosophila melanogaster"
             /mol_type="genomic DNA"
             /db_xref="taxon:7227"
             /clone="BACR08K10"
             /clone_lib="RPCI-98"
             /note="end : TET3"
ORIGIN
Query Match   6.2%; Score 81.8; DB 29; Length 1101;
Best Local Similarity 17.7%; Pred. No. 5.7e-07;
Matches 125; Conservative 300; Mismatches 278; Indels 2; Gaps 1;

Qy   538 AAAGATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGT 597
         :::|:|: ::::::|: :| :::: : | :::|:: ::: :: ::::::|: :|:::::
Db 1098 RRWGDDTWDRDTRKDDWDWTKWWTWWKDRADDRRWAGDADRWAWDDGAGTWWTATWWWWW 1039

Qy   598 ATTATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTAT 657
         :|::| | :: :::: :| :: : ::| :::::::|::::::| :::::::| |::::|:
Db 1038 WWWATWDTWWDKWWWWATAAKTDTAWTWWRTAWRADWAWGRDRGAGKRDRDAATDADGAGR 979

Qy   658 TATTATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAA 717
         :|:::: |::::|:| :::| : :|| ::::::::::|||||:|::|:|:|| |||:||
Db   978 RDGGRRKRKDKDRKDGDDDKKGGGKKKKAAKAAKWATKWWDDWDWKWDGAKDRKADD 919

Qy   718 GGTACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAA 777
         :::| : : :::::|:::|:|:::::|::::|::::::|:||::::::::::::|::::|:
Db   918 DDGAGDKDDDGKGKDADDDTDGTKDDDKDKWDDWDKAKGTWGDATWAWAATDWWWWGWA 859

Qy   778 TATGGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCCATTAGTTATTTAGAGAAG 837
         :|::::| || ::::|:||||:::|:|::|:|:|:|::::: :::::::|::::|||:|
Db   858 DADWWTWDAAADDWWADDRWDWADDRWDAWKNDDAWAWGARTADRRDWGDRAGKRGGARKRRDRK 799

Qy   838 CTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTCGTTAAACGTCGTTTC 897
         : :| :::::|:|:|::|||:|||:| ::|:::| :::::|::::|::::|||: ::::::
Db   798 RADDKRDAADDRDDRAATWTTWTTTTTIRDTDDWKWKTDTWTRWAADRTWDRDDDDDRDRAG 739

Qy   898 AATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAA 957
         ::::|||: ||:|:::||||: |::::|::::|:||:|||:|:| :|:::|:|||:|:||
Db   738 TAGRKWRRTWKRRWKRRDTRWDDADADDTARDDRRRRGGDGADAGKGKKTGRKRRRRDRA 679

Qy   958 ATTAGTGCTACAATTGAAAACAGCACGAAAGAAATATCCACATAAAGAGTTGTGCAGTA 1017
         :::::::::|:||::|||||:|::|:::||:||||||:::|:||||:|::::|::||:|
Db   678 TWDRTDAWWADAAWWTTTDTDTDDWDKRDRRRKGARRRRTTARAAWDWWTWKAWDWAKW 619

Qy 1018 TTTCAACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTA 1077
         ::::|||:|:|:::::::|| :::|:|:|:|:::::|||:|||::::|||||||:::|
Db   618 DWKTRADRWDRWAADTW--TDARKADRDWAKARAWRARRDRARARADRRWTTKGKTTTA 561

Qy 1078 AGTAAAGCAGATCGTGATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGC 1137
         |:::||::|:|:::::|:::|::|::::|||:::::||::|||||||||||||||::|
Db   560 TWTTWAARAAWWAWAWAWATTTATWTTTTTTTTTTTTTWTTTAAWWAAWWATW 501

Qy 1138 GCATTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGAT 1197
         :|:::||::|:|:|:|:::|:|::|:|||::::|||:::::::::::::::||::|||:|
Db   500 AAWTAAAWAAAAAAAAAWAATTTTTTTTTTTTWAWTAAAWTAWTWTTWTTTWWAATTT 441

Qy 1198 TCTATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTT 1242
         |::::::|::::::||||:|||||:|::|::|:::::::::|::|||
Db 1242

Qy   445 TTTGAGGCATGTGAATATAGACGTCACTTTTAAGTGTATAAACCTGATTACGCATTATG 504
          ::|| :||: :|| |: :||||| |||||
Db   306 TTGGGGGDWTTTWWWTWWTAARRRDKTTTTTTTTTTWAAAAAATATAKRGRGGADAWTAWWT 365

Qy   505 ACAAATATTGATTTCGATCATCCTGATTATTTTAAAGATATTAATGATGTTTTTGATGCA 564
          ||||| ||||| :| | ||| |||||| |||||||||
Db   366 WTWTTTTATWWWWDDDDAKAWADAGAATWTTTTTWWRAGRGARRKRDGTTDTRAADWWW 425

Qy   565 TTCCAAGAAATGGCACATAATGTTAAAAAAGGTATTATTGCTTGGGTGATGATGAACAT 624
          :|: ||::::: |||||||||||| : ::::|||||
Db   426 WTWAAWAWWWDKRKTTTTTTTTKKRRAAAAAATTTTTTTTTTTWTWAKAAAAAARRAATWT 485

Qy   625 CTACGTAAAAATTGAAGCAGCAGATGTTCCAATTATTATTATTCGGATTAAAGATTCGGATGAC 684
          :: ||||| : : : |||||||::|:
Db   486 TWWRAAAAAAAWWTWWWAARGRARGATTTTTTTAAWWNGGGAGRTAWWTWTTWTTTTTTT 545

Qy   685 ATTTATGTCAAAATATTCAAATTACGGATAAAGGTACTGCTTTTGATGTGTATGTGGAT 744
          :: ||||| : : |||||||||||||| :|||
Db   546 TTTTTTTAWWWATWWATATTTTWWTWAWAAATTTTWWTWAWAAAAASAGAWKATWTTTKKTTTATWWKRAG 605

Qy   745 GGTGAGTTTATGATCACTTCCTGTCTCCACAATGGTGACCATACACAGTTTAAATGCA 804
          |||||| ||||||| ::   ::::|||||| :|
Db   606 GATWTTTTTTTTTWWAARGDWKWTWTWTTTTTTTT------TTTWTTGTTTW 656

Qy   805 TTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTAGATGTTACACAAATATTAAAGAAGCA 864
          :::: |||||| |||:::|:|||::::|:::||||||
Db   657 WWRDGARRTAKATAKWWTWTWTWWTAWAWAGGAARKRRATWTWTWAAAWAWWAGAAAARGA 716

Qy   865 TTAGAAACGTTTGGTGGTGTTAAACGTCGTTCAATGAACTACAATTGCAAATCAAGTT 924
          :::: :|||||||| :|| |:||||||||||||||::|:|:|:||||
Db   717 GAAGRAAWTTTTTTTATKAGARRGATWTTAWWTATRARGARGADTTTTTTATAWWTAT 776

Qy   925 ATTGTAGATGATTATGCACACCATCCAAGAGAAAATTAGTGCTACA-ATTGAAACAGCACG 983
          ||||:|:|||: | |: ||||||||||| ||||
Db   777 TTTTTWWWAAGATDKAAAAAWWWWTTWTWAAAAAAATWTWWWAGARAAAKATWTWAAWAA 836

Qy   984 AAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTTCAACCACCACACTTTCTCTAGAAC 1043
          |||||| |: :||| | :|||| | ||:::
Db   837 AAAAAAAWRARWAATATATWTTTATATATAAKARAAAAATAAAAARAAARRWKGAA 896

Qy 1044 ACAGGCATTTTAAATGAATTGCAGAAGTTAAGTAAGCAGATCGTGTATTCTTATG 1103
          |||||::::|:|:: | |:||:||||| |: ||||
Db   897 AAAAAAATAAAAAWWTTWTATAAAAWWTTTWAWWRAARARAWWAAAAAAGAARAAG 956

Qy 1104 TGAAATTTTGGATCAATTGGAGAAAAATACTGGCGCATTAACGATACAAGATTTAATTGA 1163
          |||||| :: :|||| | :|||| ::: | |:
Db   957 AWAAWAATWTTTTTAATATRWWAAAAAWWTTTTTTTWTWARAAAAAWWDTTTTATWT-- 1014

Qy 1164 TAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCTATTAGAACAATTGA 1223
          ::|:|| :||:: | :::|:|:||||
Db 1015 ---ATWTAWTATAGAWAAAATARDATWRAATATWDADAARAAGAWATAKDAGAGATAAA 1070

Qy 1224 TAATGCTGTTATTTTATTTATGGTGCAGGTGATATTCAAAATTACAAAATGCATATTT 1283
          ||||||: |:|:||: | ||||||||||||| :|::|
Db 1071 WTTTTTTTTTTTTTTTTTWTAAAAAAGAAAAGAAAAATWTTARAAAAAAAARAAAAWATAAAAAAAAAA 1130

Qy 1284 AGATAAATTAGGCATGAAAAA 1304
          ||||||| ||||||:|||
Db 1131 AAAAAAAAAGAAKADRAWAA 1151

RESULT 7
CNS00EVL
LOCUS   CNS00EVL   1101 bp   DNA   linear   GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION   AL069706.1 GI:4948949
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Qy   837 GCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTT 896
          |||| :|:||||||||||||||||||||||||||
Db   789 TTTWAWAAATWACAAAAATWAAAAAAAAAWTTTAAAAAWTTWAAAAAAAWAATTTTA 848

Qy   897 CAATGAACTACAATTGCAAATCAAGTTATTGTGATGATTATGCACACCATCCAAGAGA 956
          ||| ||| | |||||| :||||||| :::|
Db   849 AAAAAAATTTTATTTTATTTATWAATTTATWAATTTATASAAAWAAAWTTTTTTTTWTT 908

Qy   957 AAATTAGTGCTACAATTGGAGAAGAAATATCCACATAAAGAAGTGTTGCAGT 1016
          :||::|| ::|: |:::|:::||||::||||||||||:|||||
Db   909 TATTTATTTWAATTWAATWTTTTTWAAWTAAAAWAWAAAAWWAAAAAAATWTWWAA 968

Qy 1017 ATTTCAACCACCACACTTTCTCTAGAACAACAGGCATTTTAAATGAATTGCAGAAAGTTT 1076
          ||||| ::|: |::|:|||||||||:|:|||::
Db   969 WWTATAAAAATANWTAWTWAAATWAAAAAWTTWAATWWWAWWTTTWAITTWWWTWTAA 1028

Qy 1077 AAGTAAAGCAGATCGTGTATTCTTATGTGAATTTTGGATCAATTAGAGAAAAATACTGG 1136
          ||||||| |:::|::| |:||||||| :|||
Db 1029 AAAAAAAATTATTTWTAWTWTTTTWTATATTTWTATATTATTTAWTAAAATWTTTWAAAWTATTW 1088

Qy 1137 CGCA---TTAACGATACAAGATTTAATTGAAGGTGCATCGTTAATTGAAGATGCTGTTAATTAATG 1192
          |||:   :::|||::::|::|:|:|:|:|||||:::|
Db 1089 ATAAAATTTTTTATTTWAWATAAAATWTAWANTWTWTWTWTTTTWTATTAAAAATAAAAATTAWATT 1148

Qy 1193 AAGATTCTATTAATGATTAGAACAATTGATAATGCTGTTATTTATTTA 1243
          ::: ||||| |||||:
Db 1149 TAWWMWATWTAAAATWTAWTTTTAWMTWAAAAAWTTTTTTTWWTTTA 1199

RESULT 6
BX437758
LOCUS   BX437758   1200 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
            5-PRIME, mRNA sequence.
ACCESSION   BX437758
VERSION   BX437758.1 GI:30773605
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1200)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope -- Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.
FEATURES   Location/Qualifiers
     source   1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP008YB01"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match   5.9%; Score 77.8; DB 13; Length 1200;
Best Local Similarity 31.6%; Pred. No. 3.9e-06;
Matches 272; Conservative 169; Mismatches 404; Indels 16; Gaps 3;

REFERENCE
AUTHORS
TITLE
JOURNAL

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT

- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match　　　　5.7%;　Score 75.4;　DB 29;　Length 1101;
Best Local Similarity　35.8%;　Pred. No. 1.3e-05;
Matches 202;　Conservative 97;　Mismatches 254;　Indels 12;　Gaps 2;

```
Qy     753 TTATGATCACTTCCTGCTCTCCACAATGGTGACCATACAGTTTAAAATGCATTAGCTGT 812
               :|| : |:::| ||| |: ::|| | ::|| | |:::|   |:::|    :: ||
Db     460 HTWTMMMMMMMWWAATWTWWAAAWAAAWTTATWAAAWAAAWAWWWWATTTTWWWWWTW 519

Qy     813 AATTGCGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTAAAGAGAGCATTAGAAAC 872
               |:| | ||| | || |:|:|| | |  :||| |  :|||| :|||
Db     520 ATTWTTWAWWTWAWTAAAAAAWATAATTAAAWWAATAWATTAAWAAATTTAWAAW 579

Qy     873 GTTTGGTGGTGTGTTAAACGTCGTTTCAATGAAACTACAAGTCAAATCAAGTTATTGTAGA 932
               :| | :|| | :|| |::|| |:|| |  || |  :||:|
Db     580 WTATATTAAWTTATAAAATWTWATTAAATAAAAAATATTTTTWTATAAAATTTTTATA 639

Qy     933 ---TGATTATGCACACCATCCAAGAGAGAATTAGTGCTACAATTGAAACAGCACGAAAGAA 989
                  || | | :| ||  |: ||| |   ||: |||:||
Db     640 AYTTAATTAWTTATTAAATAAWTWATTWWWTAAYTAAATAATTTWAAATAWAAAAAAAAA 699

Qy     990 ATATCCACATAAAGAAGTGTTGCAGTATTTCAACCACACACTTTCTCTAGAACACCAGGC 1049
               |    :||| : | |:|||   | :|
Db     700 AAAAAWATWAAWAATWATAWATAWWTTAAAAWAAAWAAAWAAWAWWATATATWW 759

Qy    1050 ATTTTTAAAATGAATTGCAGAAAGTTTAAGTAAAGCAGATCGTGTATTCTTATGTGAAAT 1109
               ||||| |:::  :|::: ::|:: ||| | :|||||| :::|:::|
Db     760 ATATATAWTTWTAWWWATWWAWWWTATAWAWWTATAWAWAAWAWAWAWAT------AAAT 810

Qy    1110 TTTTGGATCAATTAGAGAAAAATACTGGCGCATTAACGATACAAGATTTAATTGATAAAAT 1169
               ||:||  ::|| |   |    ::| |:|| ||  |
Db     811 AWATAWAWTWAAAWAWATAWWATATAWWATAWAWWAAAAAAAAATWTAATATWAATWAT 870

Qy    1170 TGAAGGTGCATCGTAATTAGAGATTCTATTAATGAGTAGAAACAATTTGATAATGC 1229
               :| | ||: :||:| : |::| | |::| ::| |:|:  ||| :|||
Db     871 AAWAAAAAAATAWAWTTWTWTTTTWAWWWATATAAAWATAAAWAWAAAWAAAAAAAAA 930

Qy    1230 TGTTATTTTATTTATGGGTGCAGTGATTATTCAAAATTCAAAAATGCATATTTAGATAA 1289
               : |:::|::|| : ::|||:|| :|: ::|:||
Db     931 TAAWAWWWTWATATWTTATTAAAWTWTAWTWAATTWAATTWWAWTWTWTATATWWTAT 990
```

───

```
Qy    1290 ATTAGGCATGAAAAATGCGTTTTAA 1314
               :|:|:| | :|::||  |:|
Db     991 ATATWWTAWTAWATATATTTATTAA 1015
```

RESULT 8
AZ548467
LOCUS　　　　AZ548467　　　908 bp　　　DNA　　　linear　GSS 14-NOV-2000
DEFINITION　ENTEK3OTR Entamoeba histolytica sheared DNA Entamoeba histolytica
genomic, genomic survey sequence.
ACCESSION　AZ548467
VERSION　　　AZ548467.1　GI:11172102
KEYWORDS　　GSS.
SOURCE　　　Entamoeba histolytica
ORGANISM　Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE　1 (bases 1 to 908)
AUTHORS　　Loftus,B., Van Aken,S. and Fraser,C.
TITLE　　　Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
JOURNAL　　Unpublished (2000)
COMMENT　　Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 828.

FEATURES
source

Location/Qualifiers
1..908
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

ORIGIN

Query Match　　　　5.7%;　Score 75;　DB 28;　Length 908;
Best Local Similarity　43.1%;　Pred. No. 1.5e-05;
Matches 363;　Conservative 0;　Mismatches 480;　Indels 0;　Gaps 0;

```
Qy     426 TGAAAGTGATTATTTCGCTTTTGAGGCATGCAATATAGACGTCACTTTTTAAGTGTTATAA 485
               || |::|:|:|:|||||| :| | |:|||| :||   :| ||||||  |||:|||
Db      33 TGAAGATGATGATGAAGAAGATGATGAGAAGAAGATGATGA 92

Qy     486 ACCTGATTACGGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAAGATAT 545
               |  |:|:|| |  ||: |:|| | :| | |:|:| |: ::
Db      93 TCGATGAAGAGATGATGAAGAAGATGATGAAGAAGAGATGATGAAGATGA 152

Qy     546 TAATGATGTTTTGATGCATTCCAAGAAATGGCACATATGTTAAAAAAGGTATTATTGC 605
               ::|:|::| :::| |: |:| : || |:::|:||  : |: |: |::| | ||
Db     153 TGAAGACGATGAAGACCGAATATGAATTAGAAGAATGATGAATGATGATGAAGAAGAAGA 212

Qy     606 TTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGCAGATGTTCCAATTTATTATTATGG 665
```

```
                source          1..1099
                                /organism="Homo sapiens"
                                /mol_type="mRNA"
                                /db_xref="taxon:9606"
                                /clone="CS0CAP002YD04"
                                /tissue_type="THYMUS"
                                /clone_lib="Homo sapiens THYMUS"
                                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                                with a NotI-oligo(dT) primer. Five prime end enriched,
                                double-strand cDNA was digested with Not I and cloned into
                                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                                Library was not normalized."

        ORIGIN

        Query Match      5.4%;    Score 71.8;  DB 13;  Length 1099;
        Best Local Similarity   26.6%;   Pred. No. 7.2e-05;
        Matches 157;  Conservative 169;  Mismatches 263;  Indels 2;  Gaps 1;

        Qy    245 AATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGA--CAGATTATTGATCAATA 302
                  ||: ||||||| ||  :|:||:|| ||:||  :::|   ::| |:| :|::
        Db   1098 AAKAAAAATKAKAKAAAWKAKAKKDAWRAADDDKATAAWKKRATKADAARADAWATKWKDD 1039

        Qy    303 TACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCACA 362
                  ||: ||::| :| |:|::|| || :: ||: |:: |||| |||| |||| |||||
        Db   1038 AAWKADAWTDAKTKKRATKRATDDKDKDKDRAAAAKADKAAKKRADTKKKDAAKAA 979

        Qy    363 TGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGTATGGGATT 422
                  |: ||:| ||||| |||||| |:| |||| |::||:  ||| ||:||| ::||
        Db    978 TKAKAKRAAKKKGKGTKDAAATKDKGTKARTDATRATDDATKDARAWWKTKKDTTK 919

        Qy    423 GCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGTTA 482
                  | ::||  |||||| | :|::|| |:|| |:| ::| |||| |:| ||::|:
        Db    918 DAKKDARAAAATAATWKAAAATWKTKKAAAWAAAAAWKAAWTKKKKKTKRWAADWAA 859

        Qy    483 TAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAAGA 542
                  ||  :::|| ||| :: |||| |:||||||||| |||| | ||:  :||:||||
        Db    858 AWAKATWDDTTKTWAADAAAAAAWRATTYKKRAWWKKARAXDAAADRWKDDKKGAKADA 799

        Qy    543 TATTAATGATGTTTTGATGCATTCCAAGAAAATGGCACATAAATGTTAAAAAGGTATTAT 602
                  |:| |||:|||||| ||||||  |||| |:| :||:| :| |||||| |||| :|
        Db    798 TKKRGTKKDKRDAKKKGGARAKKKGRAWKTKKKGAAARRAAWAARWAAAAARAAARATATK 739

        Qy    603 TGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTGTTCCAATTTATTATTA 662
                  || ::: ||| |:|:||||||| |||| ||||||  |:|| |::|||::
        Db    738 KTTKKTKKGKAATAAAATGATTGGWARKGGTKRGAAAARAARGGGGTAATRAATKGTA 679

        Qy    663 TGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTKGTTTAAATTGAAT 722
                  ::: ::::  || :|:|||| :||||||| |||| |||:|  ::|||  :|| :|||:
        Db    678 AAAAAATGRATTTGAAAAAAKRGAAWGGRAWWAAAATTKGTTTAAATTGAAAATGAAT 619

        Qy    723 TGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCAATCCTCCTGTCTCCACAATGG 782
                  ||  ||:||||| :|:||   |:| | |:|:|||| || :||: ||| |:|| |||
        Db    618 ATKTKTTTTGAWTGAATWTTTATTAATGTTTTGGTTADAWAWTTGTAAWATTGA 559

        Qy    783 TGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTATTAGA 833
                  :| :||:|||||  ::|||| ||:|| ||| |||| | :::::
        Db    558 DAAWAATARDKTTTWATAARTTAMARTKTWMAMMTHMTWAWTWTTWWWDD 508

        RESULT 10
        AZ546009/c
        LOCUS       AZ546009                 849 bp    DNA      linear   GSS 14-NOV-2000
        DEFINITION  ENTFW53TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
                    genomic, genomic survey sequence.
        ACCESSION   AZ546009
        VERSION     AZ546009.1  GI:11167130
        KEYWORDS    GSS.
        SOURCE      Entamoeba histolytica
          ORGANISM  Entamoeba histolytica
                    Eukaryota; Entamoebidae; Entamoeba.
        REFERENCE   1  (bases 1 to 849)
          AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
          TITLE     Determination of clone end sequences from Entamoeba histolytica
```

```
        Db    213 TGATGATGAAGATGATGAAGACGATGAAGACGACGAATATGAAATTAGAGATGATGATGA 272
                  ||||     |||||| |||||
        Qy    666 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATGAAAGGTACTGC 725
                       ||||
        Db    273 TGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGATGAAGATGATGAAGAAGA 332
                  ||||            ||||
        Qy    726 TTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGGTGA 785
                      ||||
        Db    333 AGAAGATGAAGAAGATGATGAAGAAGATGAAGATGATGAAGAAGATGAAGAAGAGATGA 392
                  ||||           |||||||
        Qy    786 CCATACAGTTTAAATGCATTAGCTGTAATTGCGATGACGAATATGAAATTAGAGATGATGAAGT 845
                       ||||
        Db    393 TGATGAAGATGATGAAGACGATGAAGACGACGAATATGAAATTAGAGATGATGAAGA 452
                  ||||
        Qy    846 TACAAAATATTAAAGAAGCATTAGAAACGTTTGGTGCGTGTGTTAAAACGTCGTTTCAATGAAAC 905
                      ||||
        Db    453 AGAAGAAGATGATGAAGAAGAGATGATGAAGATGATGAAGATGATGAAGAAGAAGA 512
                  ||||
        Qy    906 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 965
                      ||||
        Db    513 TGAAGATGATGAAGAAGAGATGATGAAGATGATGAAGAAGATGATGATGAAGA 572
                  ||||
        Qy    966 TACAATTGAAACAGCACGACGAAAGAAATATCCCACATAAAGAAGTTGTTGCAGTATTCAACC 1025
                       ||||
        Db    573 AGAAGATGAAGAAGATGAAGATGATGAAGAAGATGATGAAGAAGAAGATGATGATGA 632
                  ||||
        Qy   1026 ACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGTAAAGC 1085
                       ||||
        Db    633 AGACGACGAATATGAAGATGAATTAGAAGATGAATGATGAAGAAGATGAAGAAGA 692
                  ||||
        Qy   1086 AGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCATTAAC 1145
                       ||||
        Db    693 AGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGATGA 752
                  ||||
        Qy   1146 GATACAAGATTGAATAAAAATGATGAAAGTGCATCGTCATTAATGAAGAATTCTATTAA 1205
                       ||||
        Db    753 TGAAGAAGATGAAGATGATGAAGATGATGATGAAGAAGATGATGAAGATGATGAAGA 812
                  ||||
        Qy   1206 TGTATTAGACAATTGATATGCTGTTTATTTTATTTATTATGGGTGCAGGTGATATTCAAAA 1265
                        ||||
        Db    813 TGATGAAGATGATGAAGATGATGAAGATTAAAAGATGATGAAGATGATGAA 872
                  ||||
        Qy   1266 ATT 1268
                  |||
        Db    873 AAT 875

        RESULT 9
        BX455575/c
        LOCUS       BX455575              1099 bp   mRNA       linear   EST 22-MAY-2003
        DEFINITION  BX455575 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP002YD04
                    5-PRIME, mRNA sequence.
        ACCESSION   BX455575
        VERSION     BX455575.1  GI:31034783
        KEYWORDS    EST.
        SOURCE      Homo sapiens (human)
          ORGANISM  Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
        REFERENCE   1  (bases 1 to 1099)
          AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
          TITLE     Full-length cDNA libraries and normalization
          JOURNAL   Unpublished (2001)
        COMMENT     Contact: Genoscope
                    Genoscope - Centre National de Sequencage
                    BP 191 91006 EVRY cedex - France
                    Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                    Library was constructed by Life Technologies, a division of
                    Invitrogen. Contact: Feng Liang Email :fliang@lifetech.com URL :
                    http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
                    Faraday Avenue Genoscope sequence ID : CS0CAP002DB02QP1.
        FEATURES             Location/Qualifiers
```

```
JOURNAL    HM1:IMSS sheared DNA library
           Unpublished (2000)
COMMENT    Contact: Brendan J Loftus
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0208
           Fax: 301 838 3543
           Email: bloftus@tigr.org
           Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
           DNA library
           Seq primer: M13-Forward
           Class: shotgun
           High quality sequence start: 26
           High quality sequence stop: 796.
FEATURES             Location/Qualifiers
     source          1..849
                     /organism="Entamoeba histolytica"
                     /mol_type="genomic DNA"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."

ORIGIN

  Query Match             5.3%;  Score 70.4;  DB 28;  Length 849;
  Best Local Similarity   46.4%;  Pred. No. 0.00014;
  Matches    230;  Conservative   0;  Mismatches   266;  Indels    0;  Gaps    0;

Qy     520 GATCATCCTGCATTATTTTAAAGATATTAATGATGTTTTGATGCATTCCAAGAAATGGCA 579
               ||||||| |  ||    |     | | | | || |||| |||||||||||||||
Db     571 GATGATGATGAATTGAATTAGAAGATGAAGAAGATGAAGAATGATGATGAAGAAGAT 512

Qy     580 CATAAATGTTAAAAAAGGTATTATTGCTTGCGGGTGATGATGAACATCTACGTAAAATTGAA 639
               ||| |  || |    ||  ||||  | | ||||| |||||||||||||||
Db     511 GATGATGAAGATGATGATGAAGATGATGATGAAGAAGAGACGAAGAA 452

Qy     640 GCAGATGTTCCAATTTATTATTATGGATTTAAAGATTCCGATGACATTTATGCTCAAAAT 699
               || ||| |  ||    ||||||||  | |||| || || |||||||||||||
Db     451 GAAGATGATGAAGATGATTTGAATTAGAAGATGAAGAAGATGAAGAAGAT 392

Qy     700 ATTCAAAATTACCGATAAAGGTACTCGCTTTGATTGTGTATGTGGATGGTGAGTTTTATGAT 759
               |||| |||  |  | ||| |||  |||||||| |||||||||||||||||
Db     391 GATGATGATGAAGATGATGAAGATGATGAAGAAGATGATGAAGATGAAGAAGATGAT 332

Qy     760 CACTTCCTGTCTCCACAAATGGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCG 819
               ||||| | | |||| || |  | |  ||||||||||| ||||||||||||||
Db     331 GAAGATGATGAAGATGATGATGAAGATGATGATGAAGAAGATGATGATGAAGAAGAA 272

Qy     820 ATTAGTTATTAGAAGAAGCTAGATGTTACAAAATATTAAAGAAGCATTAGAAACGTTTGGT 879
               ||||| ||||||||||| | || ||||| | ||||||||||||||||||||
Db     271 TTTGAATTAGAAGATGAAGAAGATGATGATGAAGATGTGTATGTGGATGGTGAGTTTTATGAT 212

Qy     880 GGTGTTAAACGTCGTTCAATGAAATCAAGAACTACAATTGCAAATCAAGTTATTGTAGATGATTAT 939
               || ||  || | |   ||| || | || | || || |||||||||||||
Db     211 GATGAATTGAATTAGAAGATGAAGATGAAGAAGATGAAGAAGATGAAGATGAAGAAGATGAT 152

Qy     940 GCACACCATCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGGAAAGAAATATCCACAT 999
               || |  |||| || | || | |  | | || || | | ||| |||||||||||
Db     151 GAAGAAGAAGAAGAAGATGATGATGAAGAAGATGAAGAAGAAGAAGAAGATGATGAAGATGAT 92
```

```
Qy    1000 AAAGAAGTTGTTGCAG 1015
               | | || | |  ||
Db      91 GATGAAGATGATGAAG 76

RESULT 11
BJ376728/c
LOCUS       BJ376728                754 bp    mRNA    linear    EST 08-MAR-2002
DEFINITION  BJ376728 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum cDNA clone ddc29o13 3', mRNA sequence.
ACCESSION   BJ376728
VERSION     BJ376728.1  GI:19286111
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..754
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc29o13"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN

  Query Match             5.3%;  Score 69.6;  DB 12;  Length 754;
  Best Local Similarity   46.8%;  Pred. No. 0.00021;
  Matches    253;  Conservative   0;  Mismatches   285;  Indels    3;  Gaps    1;

Qy     324 TGCACATGGTAAAACTTCTACACACAGGTTTATTATCACATGTATGAATGTGATAAAAA 383
               || | | ||||  | |||  || ||| |||||||||||||||||||||||
Db     572 TGAAGAAGAGARAAAGAAATCAAAAGAAATGATGAAGAATTTGATAAATGTTAAAGAG 513

Qy     384 GACTTCATTTTAATTGGTGATGGCGCACAGGTATGGGATTGCCTGAAAGTGATTATTCGC 443
               || |  || | | ||||| | |  ||||||||||||||||||||||||||
Db     512 ATTATCAATTTTAGAAGAGAAGGAAAATAAAATGGTGATGATTATGAAGAAGAATT 453

Qy     444 TTTTGAGGCATGTGAATATGACGTCACTTTTAAGTTATTAAAACCTGATTACGCCAATTAT 503
               |||| ||| ||||| |  | || | || || |||||||||||||||||||
Db     452 TAATAAGAAATTTAATAAGAAATTAGATATTATTACAGGTAGTGATGAAGAATATGATGA 393

Qy     504 GACAAATATGATTCGATCATCCTGATTATTTTAAAGATATTAATGATGATGTTTTTGATGC 563
               || || | ||| | || |  ||| || ||||||||||||||||||||||
Db     392 TAATTATAATAACAATGATGATGAAGATGAAGATGATGATCGTGAATA 333

Qy     564 ATTCCAAGAAATGGCACCATAAATGTTAAAAAAGGTATTATTGCTTGGGGTGATGATGAACA 623
               ||||||||||| || | |||||||||||||||||||||||||||||||||
Db     332 TTATCAAGAAGAGGATTTGAAGATGAAAAAAACCTGAAAATTCAAATTATAATAAAAATAT 273

Qy     624 TCTACGTAAAATTGAAGCAGCAGATGTTCCAATTTATTATTATGGATTTAAAGATTCCGATGA 683
               | | || | ||| || | | || ||||||||||||||||||||||||||
Db     272 TGAAGAGGATGATCATGATGATGAGATGATGATTATGATGAAGGTGAGAAAAATTGTTGA 213

Qy     684 CATTTATGCTCAAAATATTCAAAATTACCGATAAAGGTACTACGTGCTTTTGATGTGTATGT--- 740
               ||||| | | || ||| | || | ||||||||||||||||||||
Db     212 ATATTATGATAAAAATGGTAATATAGTTGATATAAATGATCCAAATGTTGAATAATATTCA 153

Qy     741 GGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAATGGTGACCATACAGTTTTTAAA 800
```

```
Db    152  AGTGATGATGATAATGATGATGATAATGATGATAATGAAGGATGAAGATGAAGTGGGATGTAGAATTAAA    93
           |||  |:||  |||||                  || ||||| ||  ||  ||||||

Db    656  AAAAAMAYATHTATAWAAAAATATWTWAAWAAWAAWAAAAAAAACATWTTWCAAAAWTTWT  597

Qy    801  TGCATTAGCTGTAAATTGCCGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTAAAGA   860
           |||.|| ||  ||  |  |||| ||||||                :|  ||||||: |

Qy   1048  GCATTTTTAAATGAATTTGCAGAAAGTTAAGTAAAGCAGATCGTGTATTCTTATGTGAA  1107
           :||||:                                     :|  ||||||: |

Db     92  TGAAAGTGAATTAGAAGAAATTAAAGATTTCCATAGGATAAAAAGGGCAAGATCTTAGN   33
           |||||| ||||||||                         ||  |||||| |

Db    596  TTTTTTTTTATATATWTTTTTTTTWTTTATATMTATAAWATATATAWTTTTTTTTTW  537

Qy    861  A   861
           |

Qy   1108  ATTTTGGATCAATTAGAGAGAAAATACTGGCGCCATTAACGATACAAGATTTAATTGATAAA  1167
           ||||||||||| |  ::::||       ||| |||   |: |||:||||

Db     32  A    32
           |

Db    536  TTTTTTTTTTTTATATAWWMMMWAATWTTWTTTWAATTWTWATAAAAWTTTTTTT  477

RESULT 12
BX415058/c

Qy   1168  ATTGAAGGTGCATCGTAATTAATGAAGATTCTTAATTAATGTGTATTAGAACAATTGATAAT  1227
           :|| |:| |||||||||     |: |||||||||||:|| ||||||:|||

Db    476  TWWAAAWTTTTTTTTTTTTTTTAAATATATWYTTTTTATAAARTAWAAAAAAAMAAAWATATAT  417

LOCUS       BX415058              1056 bp    mRNA      linear    EST 15-MAY-2003
DEFINITION  BX415058 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP004YG19

Qy   1228  GCTGTTATTTTATTTATCGGTGCAGGTGATATTCAAAATTACAAAATGCATAT  1281
           :|| |  ||||||||||        |||   ||||||||

Db    416  MAWWTTTTTTTTTTTTTTTTWTATATATATATAYAAATATATATATATAT  363

3-PRIME, mRNA sequence.

ACCESSION   BX415058
VERSION     BX415058.1  GI:30767520
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1056)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP004AD10NP1.
FEATURES             Location/Qualifiers
     source          1..1056
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP004YG19"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match       5.2%;  Score 68.6;  DB 13;  Length 1056;
Best Local Similarity   34.3%;  Pred. No. 0.00034;
Matches   183;  Conservative   90;  Mismatches  258;  Indels    3;  Gaps   1;

Qy    751  TTTTATGATCACTTCCTGTCTCCACAATATGGTGACCATACAGTTTTAAATGCATTAGCT   810
           |||||::||   |||::      ::  ||| :|::        ||.:|  |

Db    896  TTWTWYAAACAAATTTWWWTTATAAAAWTTAWAAAATTTTTTTTTWWTTWWWWTWWWWW  837

Qy    811  GTAATTGCGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTATTAAGAAGCATTAGAA  870
           :: ::::          ||||||||||||||||:::::::::||||||||||||| :

Db    836  WWWWWWAWWWWAWWWWWAAAAAAAAAAAAAAHHHHHHHHHWWWWWAAAAAAAAAAAAWW  777

Qy    871  ACGTTTGGGTGGTGTAAACGTCGTTTCAATGAAACTACAATTGCAAATCAAGTTATTGTA  930
           :|||          ||||| |||    |: ||    |||||||||||

Db    776  WWWWTTWTANNNNNNNATTTTTTTTWWAAAAAAAAAWTWTTTTTTTTTTTTTTTTTTTT  717

Qy    931  GATGATTATGCACCACCATCCAAGAGAAAATTAGTGCTACAATTGAAACAGCACGAAA--G  987
           |||  |: ::|: |:|   ||: || ||||||||||:          ::

Db    716  TTTTTTTTTTWTAMAMAMTATMAMTTTTWWCTMCAYAWAWATAWTWATTWATHHTATW  657

Qy    988  AAATATCCACATAAAGAAGAGTTGTTGCAGTATTCAACCACCACACTTTCTCTAGAACACAG  1047
           ||| :|:||||  :|| ::|       || |:  |||    ||  |: |  ||
```

RESULT 13
BU498120

```
LOCUS       BU498120               594 bp    mRNA      linear    EST 11-SEP-2002
DEFINITION  PfBSToab93a09.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium
            falciparum 3D7 cDNA 5' similar to TR:Q96253 Q96253 HYPOTHETICAL
            220.3 KD PROTEIN. ; mRNA sequence.
ACCESSION   BU498120
VERSION     BU498120.1  GI:22794314
KEYWORDS    EST.
SOURCE      Plasmodium falciparum 3D7
  ORGANISM  Plasmodium falciparum 3D7
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 594)
  AUTHORS   Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
            Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
            Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
            Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
            Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU Plasmodium EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: L. David Sibley
            WashU Plasmodium EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by Debopam Chakrabarti DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: L. David Sibley
            (sibley@borcim.wustl.edu), Washington University
            Seq primer: -40UP from Gibco
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source          1..594
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="mRNA"
                     /db_xref="taxon:36329"
                     /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
                     /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
                     /note="Vector: pBluescript SK plus; Site_1: EcoRI; Site_2:
                     XhoI; Library was constructed by Debopam Chakrabarti.
                     Total RNA samples were isolated from mixed stage
                     saponin(0.1%)-lysed P. falciparum 3D7 infected
                     erythrocytes by the acidic guanidinium-phenol chloroform
                     method. The poly A+ RNA was isolated by the polyAT-Tract
                     mRNA isolation system (Promega, WI) using streptavidin
                     Magnesphere particles. Directional cDNA libraries were
                     constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
                     into EcoRI and XhoI sites of 1 ZapII vector using the Zap
                     cDNA synthesis kit (Stragene, CA). The average size of the
                     cDNA inserts in the library was between 1.0 and 1.5kb.
                     Clones were mass excised using the ExAssist helper phage
```

Left column (RESULT 14)

```
(Stratagene), the phagemids were precipitated with PEG
8000 and extracted with phenol/chloroform. Phagemid DNA
was electroporated into DH10B cells."

ORIGIN

Query Match      5.2%;   Score 68.2;   DB 13;   Length 594;
Best Local Similarity  47.0%;   Pred. No. 0.00042;
Matches 211;  Conservative  0;  Mismatches 238;   Indels  0;   Gaps  0;

Qy  491  ATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAAGATATTAATG  550
Db  123  ATAATAAAAATATTAAAGTAATGATGATGATGTTAATATATATTAAAAGTAATGATA    182

Qy  551  ATGTTTTTGATGCATTCCAAGAAATGCGCACATAAATGTTAAAAAAGTATTATTGCTTGGG  610
Db  183  ATGATAATATTAAAAGTAATGATAATTATATTAAAAGTATTATGATAATATATTAAAA    242

Qy  611  GTGATGATGAACATCTCACGTAAAATCATGAAGCAGAGATGTTCCAATTTATTATTATGGATTTA  670
Db  243  GTAATGATAATAATGATAATTAAAAATAATGATGTTAATAATAATAATAATACTTTTA    302

Qy  671  AAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGCTTTTG  730
Db  303  AAAGTGAAACTTATATTACAGTGTAAATAATAAAAACCAAAGGAGACAAAGCAGAA     362

Qy  731  ATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAGTGACCATA   790
Db  363  ATATACAAAATGGGAATACAAATAATAATAATATTATTATATATACCAATAAAAAAATAATG  422

Qy  791  CAGTTTTAAATGCATTAGCTGTAAATGCCGAATTAGTTATTAGAGAAGCTAGATGTTACAA  850
Db  423  AAAATAATAATTGCTTAAAATGATGTTTATAAATAAAAATATTTAAATGATTTAGAAATCAAG  482

Qy  851  ATATTAAAGAAGCATTAGAAAACGTTTGGTGGTGTTAAACGTCGTTTCAATGAAACTACAA  910
Db  483  AAAAGGTATTACCTAAAGGTGAAACATGATAAATGTACCAAACATATATAATGTTTTAAATA  542

Qy  911  TTGCAAATCAAGTTATTGTGATGATTAT  939
Db  543  ATACACATATATTATTATTAATTAATATAATTAT  571
```

```
RESULT 14
AL536104
LOCUS       AL536104   Homo sapiens FETAL BRAIN   1201 bp   mRNA   linear   EST 31-MAY-2003
DEFINITION  AL536104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF022YC18 5-PRIME, mRNA sequence.
ACCESSION   AL536104
VERSION     AL536104.2  GI:31260974
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:12799597.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope Sequence ID : CS0DF022BB09QP1.
FEATURES    Location/Qualifiers
     source 1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DF022YC18"
```

Right column

```
            /tissue_type="FETAL BRAIN"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL BRAIN"
            /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with Not I and
            cloned into the Not I and EcoRV sites of the pCMVSPORT 6
            vector. Library was not normalized."

ORIGIN

Query Match      5.1%;   Score 67.8;   DB 9;    Length 1201;
Best Local Similarity  32.1%;   Pred. No. 0.00049;
Matches 195;  Conservative 112;  Mismatches 298;   Indels  3;   Gaps  1;

Qy   458  AATATAGACGTCACTTTTGAAGTTATAAAACCTGATTACGCAATTATGACAAATATTGATT   517
Db   594  VAWNDVNVNVBSSKKKGGGGKGAAAAAAAAAAAAAAAAAAAAAAAAAAHAAAAAMT       653

Qy   518  TCGGATCATCCTGATTATTTTAAAGATATTAATGATGTTTTTGATGCATTCCAAGAAATGG   577
Db   654  ATATTDWWTTATATADADTDTATATAAAATTTWTATWTWTTTATTTTWTWTTTTATA       713

Qy   578  CACATAATGTTAAAAAAGGTATTATTGCTTGGGGTGATGATGAACATCTCACGTAAAAATTG  637
Db   714  WAAATATATATAATTWATTTWWTTWTWWTATATAWTTWTTWATATTTTAAATTTTTATA     773

Qy   638  AAGCAGATGTTCCAATTTATTATTATGGATTTAAAGATTCGGATGACATTTATGCTCAAA    697
Db   774  WTTTATRWWATWATWTTWTTWTTWTTWKATTTAATTAATWATWATWWTTTWTATTWTA      833

Qy   698  ATATTCAAATTACGGATAAAGGTACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATG   757
Db   834  WTAWAAAWWTTTWATATTWTTTWATATAAAATTTATTWTTTTWAATTTTATW          893

Qy   758  ATCACTTCCTGTCTCCACAATAGTGACCATACAGTTTTAAAATGTTTAAAATGCATTAGC--A  814
Db   894  TTTTTATTTATAAWTTTWTAWAWWATTWTWTWWWAAWWATTAGWTAWWAWWAWAAAATATATA  953

Qy   815  TTGCGATTAGTTATTTAGAGAAGCTAGATGTATGTTACAAATATTACAAATATTAGAAACGT  874
Db   954  WTAWGWRTAAAATAWWWWTAWAAWATWTATWTAAAAWAWAAATWWAAATATWWWTATWATT    1013

Qy   875  TTGGTGGTGTTAAACGTCGTTTCAATGAAACTACAATTGCAAATCAAGTTATTGTGATATG   934
Db  1014  WTTTWTWTTWWATWSTATWATAAAAWTAWTWTTTTTAATWTAWATAWTWTTAAAWW        1073

Qy   935  ATTATCGCACACCATCCAAGAGAAATTAGTGCTACAATTGAAACAGCACCAAAGAAATATC   994
Db  1074  AWWRTWTWTTARAAAWAWWAWTATTWRTAWWAAAAAAAATADAWATAWWWWTG          1133

Qy   995  CACATAAAGAAGTTGTTGCAGTATTCAACCACCACACTTCTCTAGACACCAGCATTTT     1054
Db  1134  ATATTATTAWATWTTWTADAWDATTTTWATATAWWWWTAGTKRTTATATATAWATATAAAAKA 1193

Qy  1055  TAAATGAA  1062
Db  1194  TAWADRWW  1201
```

```
RESULT 15
BX139987
LOCUS       BX139987   Danio rerio genomic clone DKEY-99E7   844 bp   DNA   linear   GSS 13-MAR-2003
DEFINITION  Danio rerio genomic clone DKEY-99E7, genomic survey sequence.
ACCESSION   BX139987
VERSION     BX139987.1  GI:27971314
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 844)
AUTHORS     Humphray,S.J., Huckle,E. and Durham,J.L.
```

```
TITLE       Direct Submission
JOURNAL     Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 99E7. 99E7 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1..844
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-99E7"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN
```

```
Query Match         5.1%;   Score 67.2;   DB 29;   Length 844;
Best Local Similarity  50.0%;   Pred. No. 0.00067;
Matches 168;  Conservative  0;  Mismatches 168;  Indels   0;  Gaps    0;

Qy     426 TGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGTTATAA 485
               |||||| ||||||  ||||  |||||||||||||||  |||||||||||
Db     367 TAATAATGATAATAATAATAATGATAATAATATATAATTGATAATAATGATAATAATAA 426

Qy     486 ACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCATCCTGATTATTTTAAAGATAT 545
               |||||||||| |||||||  || |||||||  || ||| |||| |||| ||||
Db     427 TAATAATAATGATAATAATAATAATATATAATTGATAATAATAATAATAATAATGATAA 486

Qy     546 TAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATTATTGC 605
               ||||||  ||||||||  ||||||||  ||||||||  ||||||||
Db     487 TAATAATATATATGATAATTGATAATAATAATAATGATAATAATAATAATAATGA 546

Qy     606 TTCGGGTGATGATCGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTATTATGG 665
               ||||||  ||||||||| |||||| |||||||||  |||||||||  |||||||
Db     547 TAATAATGATAATAATAATGATAATAATAATGATGATAATAATGATAATAATAATGA 606

Qy     666 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGC 725
               |||||||||  ||||||  || |||  ||||||||||  |||||||||||
Db     607 TGATAATAATGATAATAATAACAATAATAATAATAATGATAATAATAATAATAATGATGA 666

Qy     726 TTTTGATGTGTATGTGGATGGTGAGTTTTATGATCA 761
               ||||| |||||||||||| |||||||
Db     667 TAATAATAATAATAATAATGATGATAATAATAATAA 702
```

```
Search completed: June  6, 2004, 02:44:40
Job time : 3829 secs
```

```
                Copyright   GenCore version 5.1.6
                Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    June  3, 2004, 14:32:30 ; Search time 59 Seconds
                                     (without alignments)
                                     2092.767 Million cell updates/sec

Title:          US-10-712-713-2
Perfect score:  2275
Sequence:       1 MTHYHFVGIKGSGMSSLAQI.........GDIQKLQNAYLDKLGMKNAF 437

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq.29Jan04:*
        1: geneseqp1980s:*
        2: geneseqp1990s:*
        3: geneseqp2000s:*
        4: geneseqp2001s:*
        5: geneseqp2002s:*
        6: geneseqp2003as:*
        7: geneseqp2003bs:*
        8: geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2275	100.0	437	4	AAU00828	Aau00828 S. aureus
2	2275	100.0	444	6	ABM71210	Abm71210 Staphyloc
3	2275	100.0	444	7	ADD15465	Add15465 Staphyloc
4	2275	100.0	444	6	ADD15463	Add15463 Staphyloc
5	2274	100.0	437	6	ABU15856	Abu15856 Protein e
6	2267	99.6	437	4	AAU33801	Aau33801 Staphyloc
7	2267	99.6	444	4	AAU36700	Aau36700 Staphyloc
8	2238	98.4	437	2	AAW87771	Aaw87771 UDP-N-ace
9	1986	87.3	442	5	ABP40288	Abp40288 Staphyloc
10	1985	87.3	437	6	ABU43254	Abu43254 Protein e
11	1470	64.6	436	6	ABU18407	Abu18407 Protein e
12	1438.5	63.2	445	6	ABU29382	Abu29382 Protein e
13	1433.5	63.0	456	4	AAU35277	Aau35277 Enterococ
14	1396	61.4	444	6	ABU29727	Abu29727 Protein e
15	1342	59.0	447	5	ABB48762	Abb48762 Listeria
16	1342	59.0	447	6	ABU33064	Abu33064 Protein e
17	1298.5	57.1	443	6	ADB11616	Adb11616 Alloiococc
18	1180	51.9	457	5	ABB55394	Abb55394 Lactococc
19	1163.5	51.1	443	5	ABP27818	Abp27818 Streptoco
20	1139	50.1	442	5	ABP27819	Abp27819 Streptoco
21	1139	50.1	442	6	ABU46471	Abu46471 Protein e
22	1132.5	49.8	443	6	ABU44255	Abu44255 Protein e
23	1119	49.2	444	4	AAU37914	Aau37914 Streptoco
24	1119	49.2	444	6	ABU46149	Abu46149 Protein e
25	1117	49.1	444	6	ABU01979	Abu01979 S. pneumo
26	1067	46.9	215	2	AAW89199	Aaw89199 Partial s
27	1064	46.8	246	6	ABU43391	Abu43391 Protein e
28	1043	45.8	422	5	AAW55120	Aaw55120 Streptoco
29	1043	45.8	422	5	ABP54614	Abp54614 S. pneumo
30	1043	45.8	422	7	ADC45197	Adc45197 S. pneumo
31	967	42.5	291	7	ADC97287	Adc97287 E. faeciu
32	690.5	30.4	461	6	ABU24704	Abu24704 Protein e
33	641.5	28.2	458	6	ABU23758	Abu23758 Protein e
34	585.5	25.7	450	6	ABU25700	Abu25700 Protein e
35	574	25.2	482	6	ABU16741	Abu16741 Protein e
36	574	25.2	492	6	ADA33736	Ada33736 Acinetoba
37	559	24.6	469	6	ABU33503	Abu33503 Protein e
38	548	24.1	506	3	AAY74419	Aay74419 Neisseria
39	548	24.1	506	6	ABU38136	Abu38136 Protein e
40	545	24.0	469	6	ABU37200	Abu37200 Protein e
41	545	24.0	506	3	AAY74417	Aay74417 Neisseria
42	545	24.0	506	6	ABP80545	Abp80545 N. gonorr
43	538	23.6	468	6	ABU19357	Abu19357 Protein e
44	535	23.5	506	3	AAY74418	Aay74418 Neisseria
45	532	23.4	486	6	ABU35149	Abu35149 Protein e

ALIGNMENTS

```
RESULT 1
AAU00828
ID   AAU00828 standard; protein; 437 AA.
XX
AC   AAU00828;
XX
DT   04-JUL-2001  (first entry)
XX
DE   S. aureus UDP-N-acetylmuramate:L-alanine ligase, MurC.
XX
KW   UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
KW   antibody; wound infection; cellulitis; burn infection; eyelid infection;
KW   food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW   skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW   Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis.
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   Region          137..139
FT                   /label= Antigenic_epitope
FT   Region          236..239
FT                   /label= Antigenic_epitope
XX
PN   WO20011629 2-A2.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US023773.
XX
PR   01-SEP-1999;  99US-0151933P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Choi GH;
XX
DR   WPI; 2001-183259/18.
DR   N-PSDB; AAS00189.
XX
PT   New isolated nucleic acid for use in diagnosing Staphylococcus infections
PT   and in vaccines for eliciting immune responses to the infections.
XX
PS   Claim 9; Page 14; 225pp; English.
XX
CC   The sequence represents S. aureus MurC (UDP-N-acetylmuramate:L-alanine
CC   ligase). The polynucleotides of the invention are used to detect
CC   Staphylococcus nucleic acids in a biological sample from an animal for
CC   diagnosing Staphylococcus infections. The polypeptides of the invention
```

```
CC   are used to detect anti-Staphylococcus antibodies in a biological sample
CC   from an animal to diagnose Staphylococcus infections. The polypeptides
CC   are also used in vaccines to elicit protective antibodies in an animal to
CC   a member of the Staphylococcus genus and for preventing or attenuating an
CC   infection caused by a member of the Staphylococcus genus e.g wound
CC   infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC   joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC   scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC   disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC   polynucleotides may also be used in vaccines and for preventing or
CC   attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC   may be used to purify, detect and target the polypeptides in in vitro and
CC   in vivo diagnostic and therapeutic methods
XX
SQ   Sequence 437 AA;

Query Match             100.0%;  Score 2275;  DB 4;  Length 437;
Best Local Similarity   100.0%;  Pred. No. 9.7e-193;
Matches   437;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy     61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy    241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy    361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420

Qy    421 QKLQNAYLDKLGMKNAF 437
          |||||||||||||||||
Db    421 QKLQNAYLDKLGMKNAF 437

RESULT 2
ABM71210
ID   ABM71210 standard; protein; 444 AA.
XX
AC   ABM71210;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Staphylococcus aureus protein #450.
XX
KW   Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW   enzymatic assay; antibiotic target.
XX
OS   Staphylococcus aureus.
XX
PN   WO200294868-A2.
XX
PD   28-NOV-2002.
XX
PF   27-MAR-2002; 2002WO-IB002637.
```

```
XX
PR   27-MAR-2001; 2001GB-00007661.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V, Mora M, Scarselli M;
XX
DR   WPI; 2003-120786/11.
DR   N-PSDB; ACF72770.
XX
PT   New Staphylococcus aureus protein, useful as a vaccine for treating or
PT   preventing Staphylococcal infection, specifically an infection caused by
PT   S. aureus, e.g. sepsis.
XX
PS   Claim 1; SEQ ID NO 900; 49pp; English.
XX
CC   The invention relates to novel genes and encoded proteins from
CC   Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC   nucleic acid encoding the protein, or an antibody to the protein, is
CC   useful as a pharmaceutical, particularly as a vaccine for treating or
CC   preventing infection due to Staphylococcus bacteria, specifically an
CC   infection caused by S. aureus. The composition is particularly useful for
CC   treating or preventing sepsis in a patient. The composition can also be
CC   used for diagnostics. The protein is also used in an assay for enzymatic
CC   studies and as a target for antibiotics. This sequence represents one of
CC   the novel S. aureus proteins of the invention
XX
SQ   Sequence 444 AA;

Query Match             100.0%;  Score 2275;  DB 6;  Length 444;
Best Local Similarity   100.0%;  Pred. No. 9.9e-193;
Matches   437;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 67

Qy     61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 127

Qy    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 187

Qy    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 247

Qy    241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307

Qy    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 367

Qy    361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 427

Qy    421 QKLQNAYLDKLGMKNAF 437
          |||||||||||||||||
Db    428 QKLQNAYLDKLGMKNAF 444

RESULT 3
ADD15465
ID   ADD15465 standard; protein; 444 AA.
XX
AC   ADD15465;
XX
DT   15-JAN-2004  (first entry)
```

```
XX
DE   Staphylococcus aureus MurC protein (SeqID 30).
XX
KW   microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW   immunological response; vaccination; surface disinfectant;
KW   personal hygiene application; food preservative; enzyme; MurC;
KW   UDP-N-acetylmuramate-alanine ligase.
XX
OS   Staphylococcus aureus.
XX
PN   WO2003025007-A2.
XX
PD   27-MAR-2003.
XX
PF   20-SEP-2002; 2002WO-CA001428.
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.
PR   18-DEC-2001; 2001US-0341732P.
PR   18-DEC-2001; 2001US-0341776P.
PR   19-DEC-2001; 2001US-0341949P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A,   Dharamsi A,   Vedadi M,   Alam MZ,   Awrey D,   Beattie B;
PI   Domagala M,   Houston S,   Kanagarajah D,   Nethery K,   Ng I,   Mansoury K;
PI   Mcdonald M,   Pinder B,   Viola C,   Wrezel O;
XX
DR   WPI; 2003-468119/44.
DR   N-PSDB; ADD15464.
XX
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 30; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polypeptide sequence
CC   is the MurC protein encoded by experimentally predicted DNA from S.
CC   aureus UDP-N-acetylmuramate-alanine ligase of the invention.
XX
SQ   Sequence 444 AA;

Query Match        100.0%;    Score 2275;   DB 7;   Length 444;
Best Local Similarity    100.0%;    Pred. No. 9.9e-193;
Matches 437;    Conservative 0;    Mismatches 0;    Indels 0;    Gaps 0;

Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  67

Qy   61  DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
```

```
Db   68  DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 127

Qy  121  HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
Db  128  HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 187

Qy  181  DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
Db  188  DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 247

Qy  241  TAFDVYVDGBFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Db  248  TAFDVYVDGBFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307

Qy  301  ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
Db  308  ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 367

Qy  361  KADRVFLCEIFPGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
Db  368  KADRVFLCEIFPGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 427

Qy  421  QKLQNAYLDKLGMKNAF 437
Db  428  QKLQNAYLDKLGMKNAF 444

RESULT 4
ADD15463
ID   ADD15463 standard; protein; 444 AA.
XX
AC   ADD15463;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Staphylococcus aureus MurC protein (SeqID 28).
XX
KW   microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW   immunological response; vaccination; surface disinfectant;
KW   personal hygiene application; food preservative; enzyme; MurC;
KW   UDP-N-acetylmuramate-alanine ligase.
XX
OS   Staphylococcus aureus.
XX
PN   WO2003025007-A2.
XX
PD   27-MAR-2003.
XX
PF   20-SEP-2002; 2002WO-CA001428.
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.
PR   18-DEC-2001; 2001US-0341732P.
PR   18-DEC-2001; 2001US-0341776P.
PR   19-DEC-2001; 2001US-0341949P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A,   Dharamsi A,   Vedadi M,   Alam MZ,   Awrey D,   Beattie B;
PI   Domagala M,   Houston S,   Kanagarajah D,   Nethery K,   Ng I,   Mansoury K;
PI   Mcdonald M,   Pinder B,   Viola C,   Wrezel O;
XX
DR   WPI; 2003-468119/44.
DR   N-PSDB; ADD15462.
XX
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
```

PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 28; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polypeptide sequence
CC   is the MurC protein encoded by DNA predicted from the genomic sequence of
CC   S. aureus UDP-N-acetylmuramate-alanine ligase of the invention.
XX
SQ   Sequence 444 AA;

Query Match   100.0%; Score 2275; DB 7; Length 444;
Best Local Similarity   100.0%; Pred. No. 9.9e-193;
Matches   437;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 67

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 127

Qy   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACYRRHFLSYKPDYAIMTNIDFPDHPDYFK 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACYRRHFLSYKPDYAIMTNIDFPDHPDYFK 187

Qy   181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKG 247

Qy   241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307

Qy   301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 367

Qy   361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 427

Qy   421 QKLQNAYLDKLGMKNAF 437
              |||||||||||||||||
Db   428 QKLQNAYLDKLGMKNAF 444

RESULT 5
ABU15856
ID   ABU15856 standard; protein; 437 AA.
XX
AC   ABU15856;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #1383.
XX

KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO20277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD, Carr GJ,   Yamamoto R,   Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA19726.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 43780; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 437 AA;

Query Match   100.0%; Score 2274; DB 6; Length 437;
Best Local Similarity   99.8%;   Pred. No. 1.2e-192;
Matches   436;   Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 437 AA;

Query Match   99.6%;   Score 2267;   DB 4;   Length 437;
Best Local Similarity 99.5%; Pred. No. 5e-192;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 MTTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILLPFDANNIKE  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    1 MTTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILLPFGANNIKE  60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDARQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDARQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437
```

RESULT 7
AAU36700
ID   AAU36700 standard; protein; 444 AA.
XX
AC   AAU36700;
XX
DT   14-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #870.
XX
KW   Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.

```
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDARQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDARQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437
```

RESULT 6
AAU33801
ID   AAU33801 standard; protein; 437 AA.
XX
AC   AAU33801;
XX
DT   14-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #77.
XX
KW   Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS51660.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Example 3; SEQ ID NO 5297; 51pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes

```
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVLFMGAGDI 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVLFMGAGDI 427

Qy  421 QKLQNAYLDKLGMKNAF 437
            |||||||||||||||||
Db  428 QKLQNAYLDKLGMKNAF 444
```

RESULT 8
AAW87771
ID   AAW87771 standard; protein; 437 AA.
XX
AC   AAW87771;
XX
DT   17-MAR-1999 (first entry)
XX
DE   UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide).
XX
KW   MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW   bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW   immunogen; drug; genetic immunisation.
XX
OS   Staphylococcus aureus.
XX
PN   EP889123-A2.
XX
PD   07-JAN-1999.
XX
PF   26-JUN-1998;   98EP-00305064.
XX
PR   03-JUL-1997;   97US-0052720P.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
PA   (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI   Wallis NG, Burnham MKR;
XX
DR   WPI; 1999-062655/06.
DR   N-PSDB; AAV99650.
XX
PT   New isolated MurC polypeptide from Staphylococcus aureus and related
PT   nucleic acid - useful in diagnosis, treatment and prevention of bacterial
PT   infections.
XX
PS   Claim 1; Page 4; 39pp; English.
XX
CC   The present sequence represents a UDP-N-acetylmuramate:L-alanine ligase
CC   (MurC polypeptide) encoded by the S. aureus MurC gene. Host cells
CC   containing an expression system comprising the MurC gene can be used for
CC   the recombinant production of the polypeptide. Agonists or the MurC
CC   polypeptide are used to treat conditions requiring increased activity or
CC   expression of the polypeptide. Antagonists, inhibitory nucleic acid or
CC   competitive polypeptide are useful for inhibiting the polypeptide e.g.
CC   bacterial (especially S. aureus) infections. They are also useful against
CC   Helicobacter pylori infections and related cancers, ulcers and gastritis.
CC   The antibacterial agents are useful to treat in-dwelling devices for
CC   infection prevention or generally as wound treatments to prevent adhesion
CC   of bacteria to matrix proteins. The MurC polypeptide is also useful for
CC   diagnosing or prognosing a (susceptibility to) disease, for raising
CC   antibodies; to identify modulators or specific receptors; in rational
CC   drug design and as an immunogen for vaccines. The MurC gene sequences are
CC   useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
CC   for chromosomal mapping; to determine bacterial serotype; and for genetic
CC   immunisation
XX
SQ   Sequence 437 AA;

Query Match   98.4%; Score 2238; DB 2; Length 437;
Best Local Similarity   98.2%; Pred. No. 1.8e-189;
Matches 429; Conservative   3; Mismatches   5; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
```

us-10-712-713-2.rag

XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS54559.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Example 3; SEQ ID NO 12293; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence represents an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 444 AA;

Query Match             99.6%;  Score 2267;  DB 4;  Length 444;
Best Local Similarity   99.5%;  Pred. No. 5.1e-192;
Matches  435;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

```
Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFGANNIKE 67

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 127

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 187

Qy  181 DINDVFDAFQEMAHNVKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 DINDVFDAFQEMAHNVKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 247

Qy  241 TAFDVVYDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TAFDVVYDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 367
```

```
Db         |||||||||||||||||||||||||||||||||||||| ||||||||
     1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPLPFGANNIKE  60
Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS  120
Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK  180
Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG  240
Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN  300
Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKSVVAVFQPHTFSRTQAFLNEFAESLS  360
        |||||||||||||||||||||||| ||||||||||:||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIDTARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLC  360
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI  420
        |||||||||||||||||:|||||||||||| ||| ||||:|||||||||||:||||||||
Db  361 KADRVFLCEIFGSIRENSGALTIQDLIDKIGGASFINEDLINVLEQFDNAVVLFMGAGDI  420
Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437
```

```
RESULT 9
ABP40288
ID   ABP40288 standard; protein; 442 AA.
XX
AC   ABP40288;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5133.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-00134001.
XX
PR   14-AUG-1997;   97US-0055779P.
PR   08-NOV-1997;   97US-0064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN92833.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT   polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS   Disclosure; SEQ ID NO 5133; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences can
CC   also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life cycle
CC   or inhibit S. epidermidis infection. N.B. The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site
XX
SQ   Sequence 442 AA;
```

```
Query Match             87.3%;  Score 1986;  DB 5;  Length 442;
Best Local Similarity   84.7%;  Pred. No. 4e-167;
Matches  370;  Conservative  43;  Mismatches  24;  Indels   0;  Gaps    0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPLPFDANNIKE  60
        |||||||||||||||||||||||||||||||||:||||||||||||||||| :||||
Db    6 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIESYVFTEVALRNKGIKILPFDANNITK  65
Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS  120
        :||||||||:|||||||||||||||| ::  :|||| :||||||||||||||||||||||
Db   66 EMVVIQGNAFPDNHEEIVRAHELKLDIIKYHDFLGHVINQYTSVAVTGAHGKTSTTGLLS  125
Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK  180
        |||||||||||||||||||||| ||||||||||||||||||:|||||||||||||||||
Db  126 HVMNGDKKTSFLIGDGTGMGLPGSDYFAFEACEYRRHFLSYHPDYAIMTNIDFDHPDYFK  185
Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG  240
        :|:||:||  |||:|||||||||||||:||:|::||:|||:||::||||| ::||||:||
Db  186 NIDDVYDAFQHMALNVKKGIIAWGDDRYLRKLDVDIPVYYYGFKETDDIYAKNIQITEKG  245
Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN  300
        | ||||: ||||| |||||||| ||:||||||||:||:: :: :||||: ||||||||||
Db  246 TQFDVYIKGEFYDQFLSPQYGNHNLNALAVIAISYLEMNVENIKEALITFGGVKRRFN  305
Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKSVVAVFQPHTFSRTQAFLNEFAESLS  360
        ||:::  |||||||||||||||||||||||: :||||||||||||||||||||||||||
Db  306 ETKVSNQVIVDDYAHHPREISATIETARKKYPQKDVVAVFQPHTFSRTQAFLNEFAESLS  365
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI  420
        |||:|||||||||||||||:||:|||:|:||::|:||||:|:|||:|||||:||||||||
Db  366 KADQVFLCEIFGSIRENTGDLTIEDLINRIDGSTLIDENSIDVLEKFDNAVILFMGAGDI  425
Qy  421 QKLQNAYLDKLGMKNAF 437
        ||| :||::|||:||:|
Db  426 QKLLKAYFEKLGVKNDF 442
```

```
RESULT 10
ABU43254
ID   ABU43254 standard; protein; 437 AA.
XX
AC   ABU43254;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #28781.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002;  2002WO-US009107.
XX
PR   21-MAR-2001;  2001US-00815242.
PR   06-SEP-2001;  2001US-00948993.
PR   25-OCT-2001;  2001US-0342923P.
PR   08-FEB-2002;  2002US-0007285L.
PR   06-MAR-2002;  2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
```

```
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA47124.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 71178; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence data for this
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 437 AA;

Query Match      87.3%;  Score 1985;  DB 6;  Length 437;
Best Local Similarity   84.7%;  Pred. No. 4.9e-167;
Matches 370;  Conservative 43;  Mismatches 24;  Indels 0;  Gaps 0;

Qy    1 MTHYHFVGIKGSGMSSLLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||| ;
Db    1 MTHYHFVGIKGSGMSSLLAQIMHDLGHEVQGSDIESYVFTEVALRNKGIKILPFDANNITK 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          :|||||| :| |||| :|:|||||| |||  |||   :|:|||||||||||||||||||||
Db   61 EMVVIQGNAFPDNHEEIVRAHELKLDIIKYHDFLGHVINQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          |||||||||||||||||||||| |||||| |||||||||||:||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPGSDYFAFEACEYRRHFLSYHPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVYFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKOSDDIYAQNIQTDKG 240
          :|:|||  ||| ||:|||||||||||||:|||:  |:|:||||||  |||||:|||| |||
Db  181 NIDDVYDAFQHMALNVKKGIIAWGDDEYLRKLDVDIPVYYYGFKEFDDIYAKNIQTEKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          | |||| . ||||  ||||||:|  :||||||||||||||:  : |||::||||||||||||
Db  241 TQFDVYIKGEFYDYQFLSPQYGNHNILNALAVIAISYLEDMVENIKRALITFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          || :.|||||||||||||||||||||||||||:|:||||||||||||||||||||||:||
Db  301 ETKVSNQVIVDDYAHHPREISATIETARKKYPQKDVVAVFQPHTFSRTQAFLNEFABSLS 360
```

```
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          |||:|||||||||||||||:|||:|||:|| |:.||:|::|:||:.|||||||||||||||
Db  361 KADQVFLCEIFGSIRENTGDLTIEDLINRIDGSTLIDENSIDVLEKFDNAVILFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
          ||| |||:.|||:||  |
Db  421 QKLLKAYFEKLGVKNDF 437

RESULT 11
ABU18407
ID   ABU18407 standard; protein; 436 AA.
XX
AC   ABU18407;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #3934.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Bacillus anthracis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-0081524 2.
PR   06-SEP-2001; 2001US-0948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-0007285 1.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (BLIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA22277.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 46331; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
```

CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 436 AA;

Query Match   64.6%; Score 1470; DB 6; Length 436;
Best Local Similarity   64.7%; Pred. No. 2.1e-121;
Matches   279; Conservative 59; Mismatches 93; Indels   0; Gaps   0;

```
Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
        ||| |||||||||||||||||||:|||||||:|:  .||:  :|:|
Db    1 MTVYHFVGIKGTGMSSLAQILHDMKHTVQGSDYEKRFFTQTALEKRNISILPFDKSNVKE 60

Qy   61 DMVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||  |:  ||||: |: ||:  :||||||||||||||||
Db   61 GQVIAGNAFPDTHEEIVAAKELNIPVHRYHHFLGDLMNQYTSVAVTGAHGKTSTTGLLA 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||  :||||||||||| :     |
Db  121 HVMQGAHPTSYLIGDGTGHGVENSKYFVFEACEYRRHFLSYNPDYAIMTNIDFDHPDYFT 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIBADVPIYYYGFKDSDDIYAQNIQITDKG 240
        ||||| |||||| || ::|:  |:  |:|||
Db  181 DINDVFSAFQEMALQVKFGIACGDDEELQKIQAKVPVIFYGFGEDNDFQARNIQKRTDG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        ||||| |||:|   |||
Db  241 TIFDVFVRNTYYDTFKITGYGNHSVLNALAVIALCHYENVDVEAVKQLTTFEGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATITARKKYPHKBVVAVFQPHTFSRTQAFLNEFABSLS 360
        :||||:  ||||| |
Db  301 EKPMGEQVIIDDYAHHPTENATIEAARQKHPEREIVAVFQPHTFSRTBKFLDEFABSLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQPDNAVILFMGAGDI 420
        ||:  ||||:  |||
Db  361 KADQVYLCDIFGSARENKGELTIEDLQKRIDGAELITDTTTDVLKKGHKNGVLIFMGAGDI 420

Qy  421 QKLQNAYLDKL 431
        ||| |:  ::
Db  421 QKFEAAYVKEV 431
```

RESULT 12
ABU29382
ID   ABU29382 standard; protein; 445 AA.
XX
AC   ABU29382;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #14909.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.

XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA33252.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 57306; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 445 AA;

Query Match   63.2%; Score 1438.5; DB 6; Length 445;
Best Local Similarity   60.9%; Pred. No. 1.4e-118;
Matches   259; Conservative 79; Mismatches 86; Indels   1; Gaps   1;

```
Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
        |||||||||||||||:  :|.| :| :|||||: ||||: :||||||||:|||:
Db    9 YHFVGIKGSGMSSLALVLHQKGYNVQGSDVEEYFFQRDLEKSGVPILPFNADNIDKDMI 68

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
        ||:|||| :  :||||: :| :  ||||:|||: ||||: ||||||:||||||||:|:
Db   69 VIAGNAFPDTHEBIARAIELGABVIRYHDFIARFIEPYTSIAVTGSHGKTSTTGLIAHVL 128

Qy  124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
        ||:  :||||||||:|||   :  :|||||||||||:||||||||||||||||||: |
Db  129 SGINPTSYLIGDGTGHGEPDADFFAFEACEYRRHFLAYSPDYAIMTNIDFDHPDYYKSIE 188

Qy  184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIRADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
        |||: |||:||| || :  ||||: :||  ||:||||||  ::|||
Db  189 DVFSAFQTMAHQVKKGIFAYGDDKYLRQLBSEVPVYYYGVSEBDDIQARNIORTTEGSSF 248

Qy  244 DVYVDGEFYDHFLSPQYGDHTVLNALAVIAVYFEKLDMQKVAEEMLSFKGVKRRFSEKK 303
        |||   ||| :|  |||:|||  ::| |||||||| :|
Db  249 DVYHKDDFVGHFVLPARGHHNIMNALGVIAVAYFEKLDMQKVAEEMLSFKGVKRRFSEKK 308
```

```
XX
SQ   Sequence 456 AA;
     Query Match        63.0%;  Score 1433.5;  DB 4;  Length 456;
     Best Local Similarity  60.7%;  Pred. No. 3.9e-118;
     Matches 258;  Conservative 79;  Mismatches 87;  Indels 1;  Gaps 1;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
         |||||||||||||||| :||:: :||||:: ||:: :||:|:|:: |: :|||||:|
Db   20 YHFVGIKGSGMSSLALVLHQKGYNVQGSDVEEYYFTQRDLEKSGVPILPFNADNIDKDMI 79

Qy   64 VIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
         :|| ||| :|||| : :|: : :|:||||:|||||||||||||||||:|||||||||:|
Db   80 VIAGNAFPDTHEIARAIELGAEVIRYHDFIARPIEPYTSIAVTGSHGKTSTTGLLAHVL 139

Qy  124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
         .: .||::|||||||.|| .:|:.|||||||||||||:::|||||||||||||||:|:.|
Db  140 SGINPTSYLIGDGTGHGEPDADFFAFEACEYRRHFLAYSPDYAIMTNIDFDHPDYKSIE 199

Qy  184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
         |||.|||.||||:|||:||:|||:||||:|:||||||| |::.||||||||||:|::::|
Db  200 DVFSAFQTMAHQVKKGIFAYGDDKYLRQLESEVPYYYQGVSEEDIQARNIQRTEGSSF 259

Qy  244 DVVVDGEPYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETPGGVKRRPNETT 303
         || .. :|.||||.|..|||:::||| ||| ||:|||||::|:|:::::|||||:|::..
Db  260 DVYHKADFVGHFVLPAFGHHNIMNALGVIAVAYFEKLDMQKVABMLSFKGVKRRFSEKK 319

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
         :::. :|||||||||.| :.|.:.|||||||||:|||||||||:||.|:|:||||:|||.|
Db  320 VSDMIIVDDYAHHPABIKATIDGARQKYPDKEIIAVFQPHTFTRTIALMDEFAEALDLAD 379

Qy  364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQFDNAVILFMGAGDIQK 422
         .|||||:|||.|||.| :|::|||:||:|  || :||:::|||||::||:|||||||||:||
Db  380 EVFLCNIFGSARETQGEVRIEDLGEKIQKGGQVITEDNVSPLLDFENAEVVFMGAGDVQK 439

Qy  423 LQNAY 427
         :||
Db  440 FEQAY 444


RESULT 14
ABU29727
ID   ABU29727 standard; protein; 444 AA.
XX
AC   ABU29727;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #15254.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Enterococcus faecium.
XX
PN   WO20027183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-0072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW,
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA33597.
```

```
Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
         :::. :|||||||||.| :.|.:.|||||||||:|||||||||:||.|:|:||||:|||.|
Db  309 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFTRTIALMDEFAEALDLAD 368

Qy  364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQFDNAVILFMGAGDIQK 422
         .|||||:|||.|||.| :|::|||:||:|  || :||:::|||||::||:|||||||||:||
Db  369 EVFLCNIFGSARETQGEVRIEDLGEKIQKGGQVITEDNVSPLLDFENAVVFMGAGDVQK 428

Qy  423 LQNAY 427
         :||
Db  429 FEQAY 433


RESULT 13
AAU35277
ID   AAU35277 standard; protein; 456 AA.
XX
AC   AAU35277;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis cellular proliferation protein #564.
XX
KW   Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW   antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,   Ohlsen KL,   Zyskind JW,   Wall D,   Trawick JD,   Carr GJ;
PI   Yamamoto RT,   Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS53136.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Example 3; SEQ ID NO 10870; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence represents an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
```

XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 57651; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 444 AA;

Query Match   61.4%; Score 1396; DB 6; Length 444;
Best Local Similarity   60.6%; Pred. No. 7.9e-115;
Matches 260; Conservative 74; Mismatches 93; Indels 2; Gaps 2;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKBDMV 63
           ||||||||||||||||| :||: ||||||| ||:: |||| | ||:|:|
Db   8 YHFVGIKGSGMSSLALVLHEQGLNVQGSDIEKYFFQRDLEKANITILPFNADNVKPGMT 67

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
           ||||| | || ||||||||||||:||:|||:||||||| :|||||||||||||||:||
Db   68 IIAGNAFPDSHEEIQRAKELGLEVIRXHDFIGHFIQNYTSIAVTGSHGKTSTTGLLSHVL 127

Qy   124 NGDKKTSFLIGDTGCMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFHPDYFKDIN 183
           :|:: ||||||||||||||||||||||| ||||||||||||
Db   128 SGVRPTSYLLGDTGCHGDPQABFFSPEACEYRRHFLAYSPDYAIMTNIDFHPDYTTSID 187

Qy   184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKGTAP 243
           ||||||:| :||||||||:| |||:||: |||||| :: ||| :| ||:|||:
Db   188 DVFTAFQTMAGQVKKAIFAYGDDAYLRKLKANVPIYYGVTENDDIQARNIERTTSGSAF 247

Qy   244 DVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
           |||:|| |||| |||| : || |:|||||||:||||| :|||||:
Db   248 DVYHGDEPVGHFTVPAFGKHNILNALGVIAVAYFEKLDLKEVAEEMLTPFGVKRRFSEKI 307

Qy   304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
            :|: |||||||||:| ::|||||| ||:|| ||||||||||:|||
Db   308 VADMTVVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFTRTIALMDEFAEALDLAD 367

Qy   364 RVFLCBIFGSIRENTGALTQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDIQK 422
           :|:|||||:|| | :| |: |||::: |||: |:|||||||||||||||
Db   368 KVYLCDIFGSAREEQNVKIEDLGAKIKKGGEVIKENNVSPLLDYHDAVVIFMGAGDVQK 427

Qy   423 LQNAYLDKL 431
            :|| ||| :||
Db   428 FEQAY-EKL 435

RESULT 15
ABB48762
ID ABB48762 standard; protein; 447 AA.
XX
AC ABB48762;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1466.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR001118.
XX
PR 11-APR-2000; 2000FR-00004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and related
PT polypeptides.
XX
PS Claim 6; SEQ ID NO 1467; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 447 AA;

Query Match   59.0%; Score 1342; DB 5; Length 447;
Best Local Similarity   59.7%; Pred. No. 4.9e-110;
Matches 255; Conservative 64; Mismatches 108; Indels 0; Gaps 0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

```
Db        |||||||||||||||||||||||| :|||||||:: :|||||:: ||: || ||: |:|||:|
     1 MTIYHFVGIKGSGMSALAQILHDKGFQVQGSDVDKYFFTQKALEEKQIPIMTFSADNIQE  60

Qy    61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          : :|||| :|||| |: |: :|| |:  ||: |: |: ||:|:||||||||:||||||
Db    61 GLTIIAGNAFPDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKTSTTGLLS 120

Qy   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||: : ||||||:||||||||  :: : |: ||:|||: ||||||||||||||:|||||
Db   121 HVVGAIRPTSYLIGDGTGSGTKDAEYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK 180

Qy   181 DINDVFDAFQEMAHNVKKGIIAWGDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ::||:|||:: :|||: : |||: : |::| |||:||:: |||:: :|:|:
Db   181 SVDDVFNAFETLGKQVKKAVFALGDDAELRKLTLDIPIIYFGFGEENEFQAKNVIKETTG 240

Qy   241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          | ||||| |||: |||||:| ||  ||| |: : ||||| :|||||||||
Db   241 TKFDVYHREEFLSSFEIPAYGDHNVLNALSVIALCDYEGLPVEDVKNELKTFEGVKRRFS 300

Qy   301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQARLNEFAESLS 360
          |  : |||:|||||||||||: | |: |::||: :|||| ||| |::|: :|:|:
Db   301 ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPDKKVVAVFQPHTFTRTRTFLQGFADSLN 360

Qy   361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          || |:|:|:||| ||| ||  :: :|| ||||||||||||||
Db   361 LADEVYLCDIFGSAREKTGNLIIADLIAHKTKGNHIIKEEHTEELLKYPEAVILFMGAGDV 420

Qy   421 QKLQNAY 427
          || |:||
Db   421 QKFQAAY 427
```

Search completed: June 3, 2004, 14:39:09
Job time : 61 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   June  3, 2004, 14:37:01 ; Search time 22 Seconds
                                     (without alignments)
                                     1025.480 Million cell updates/sec

Title:        US-10-712-713-2
Perfect score: 2275
Sequence:     1 MTHYHFVGIKGSGMSSLAQI.........GDIQKLQNAYLDKLGMKNAF 437

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:     389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued Patents AA:*
            1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
            2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
            3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
            4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
            5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
            6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1986	87.3	442	4	US-09-134-001C-5133	Sequence 5133, Ap
2	1119	49.2	444	4	US-08-940-572-2	Sequence 2, Appli
3	1043	45.8	422	3	US-08-961-083-116	Sequence 116, App
4	1043	45.8	422	4	US-09-536-784-116	Sequence 116, App
5	967	42.5	291	4	US-09-107-532A-6914	Sequence 6914, Ap
6	574	25.2	492	4	US-09-540-236-5023	Sequence 5023, Ap
7	532	23.4	493	4	US-09-540-236-2712	Sequence 2712, Ap
8	499	21.9	492	4	US-09-489-039A-12280	Sequence 12280, A
9	490	21.5	513	4	US-09-252-991A-24196	Sequence 24196, A
10	475	20.9	488	4	US-09-543-681A-6951	Sequence 6951, Ap
11	470	20.7	812	4	US-09-198-452A-978	Sequence 978, App
12	417	18.3	475	4	US-09-328-352-4942	Sequence 4942, Ap
13	386.5	17.0	473	4	US-09-540-236-2891	Sequence 2891, Ap
14	371	16.3	471	4	US-09-543-681A-5705	Sequence 5705, Ap
15	358.5	15.8	458	4	US-09-489-039A-8987	Sequence 8987, Ap
16	265	11.6	494	3	US-08-984-618-3	Sequence 3, Appli
17	232.5	10.2	97	4	US-08-940-572-4	Sequence 4, Appli
18	199	8.7	457	4	US-09-134-001C-3838	Sequence 3838, Ap
19	194	8.5	283	4	US-09-252-991A-21952	Sequence 21952, A
20	194	8.5	449	1	US-09-530-836-2	Sequence 2, Appli
21	187.5	8.2	450	1	US-08-665-435A-2	Sequence 2, Appli
22	187.5	8.2	450	2	US-08-843-309-2	Sequence 2, Appli
23	185	8.1	46	4	US-08-936-165A-446	Sequence 446, App
24	183.5	8.1	446	2	US-08-934-481-2	Sequence 2, Appli
25	183.5	8.1	446	4	US-09-290-602-2	Sequence 2, Appli
26	180	7.9	413	3	US-09-147-928-2	Sequence 2, Appli
27	177	7.8	457	4	US-09-134-001C-4284	Sequence 4284, Ap
28	175.5	7.7	455	4	US-09-107-532A-5228	Sequence 5228, Ap
29	164.5	7.2	452	4	US-09-530-836-6	Sequence 6, Appli
30	160	7.0	335	3	US-08-961-083-110	Sequence 110, App
31	160	7.0	335	4	US-09-536-784-110	Sequence 110, App
32	160	7.0	451	4	US-09-530-836-5	Sequence 5, Appli
33	159	7.0	452	4	US-09-144-918-2	Sequence 2, Appli
34	156	6.9	483	4	US-09-198-452A-438	Sequence 438, App
35	155	6.8	437	4	US-09-530-836-4	Sequence 4, Appli
36	152	6.7	478	4	US-09-489-039A-12376	Sequence 12376, A
37	151	6.6	419	4	US-09-198-452A-974	Sequence 974, App
38	150.5	6.6	523	4	US-09-540-236-3139	Sequence 3139, Ap
39	146	6.4	488	3	US-08-984-618-4	Sequence 4, Appli
40	144	6.3	483	4	US-09-328-352-6020	Sequence 6020, Ap
41	139.5	6.1	502	4	US-09-543-681-7027	Sequence 7027, Ap
42	135	5.9	469	4	US-09-540-236-2593	Sequence 2593, Ap
43	134.5	5.9	465	4	US-09-107-532A-7268	Sequence 7268, Ap
44	132.5	5.8	448	3	US-09-143-954-4	Sequence 4, Appli
45	132.5	5.8	457	3	US-09-143-954-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-134-001C-5133
; Sequence 5133, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5133
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5133

Query Match             87.3%;  Score 1986;  DB 4;  Length 442;
Best Local Similarity   84.7%;  Pred. No. 1.1e-185;
Matches 370;  Conservative  43;  Mismatches  24;  Indels   0;  Gaps   0;

Qy      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVFTEVALRNKGIKILPFDANNIKE 60
          ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||:
Db      6 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIESVFTEVALRNKGIKILPFDANNITK 65

Qy     61 DMVVIQGNARASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          :|||||||| :.|||||||:.||.:|.|: |.|||.:|:|||||||||||||||||||||
Db     66 EMVVIQGNAFPDNHEBIVRAHELKLGLDIIKYHDFLGHVINQYTSVAVTGAHGKTSTTGLLS 125

Qy    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||||||||||||||||||||||.|||||||||||||||||||:||||||||||||||||||
Db    126 HVMNGDKKTSFLIGDGTGMGLPGSDYFAFEACEYRRHFLSYHPDYAIMTNIDFDHPDYFK 185

Qy    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKOSDDIYAQNIQITDKG 240
          :|:|:||||||:|.||||||||||||::|||::|::||:|||||.::|||||:||||:||
Db    186 NIDDVYDAFQHMALNVKKGIIAWGDDEYLRKLDVDIPVYYYGFKETDDIYAKNIQITEKG 245

Qy    241 TAFDYVIKGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          |.||:.|||||||.|||||||.|..|||||||||||||..:.:.:.|||..||||||||||
Db    246 TQFDVYIKGEFYDQFLSPQYGNHNILNALAVIAISYLENMNVENIKEALITFGGVKRRFN 305

Qy    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAFQPHTFSRTQAFLNEFAESLS 360
          || ::.||||||||||||||||||||||||||.|:.||.|||||||||||||||||||||
Db    306 ETKVSNQVIVDDYAHHPREISATIETARKKYPQKDVVAVFQPHTFSRTQAFLNEFAESLS 365
```

```
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        ||:||||||||||||||:  :|||:|:::||||||:||:|||| :|||||||||||||||||
Db  366 KADQVFLCEIFGSIRENTGDLTIEDLLINRIDGSTLIDENSIDVLEKFDNAVILFMGAGDI 425

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||: ||::|||:||:|
Db  426 QKLLKAYFEKLGVKNDF 442
```

RESULT 2
US-08-940-572-2
; Sequence 2, Application US/08940572
; Patent No. 6310193
; GENERAL INFORMATION:
;   APPLICANT: Wallis, Nicola G.
;   APPLICANT: Black, Michael T.
;   APPLICANT: Hodgson, John E.
;   APPLICANT: Knowles, David J.
;   APPLICANT: Lonetto, Michael A.
;   APPLICANT: Nicholas, Richard O.
;   APPLICANT: Stodola, Robert K.
;   TITLE OF INVENTION: No. 6310193el MurC
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dechert, Price & Rhoads
;   STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;   CITY: Philadelphia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19103-2793
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/940,572
;   FILING DATE:
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/024022
;   FILING DATE: 16-AUG-1996
;   APPLICATION NUMBER: US 08/889711
;   FILING DATE: 08-JUL-1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Dickinson, Q. Todd
;   REGISTRATION NUMBER: 28,354
;   REFERENCE/DOCKET NUMBER: P50533-04
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 215/994-2252
;   TELEFAX: 215/994-2222
;   TELEX:
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 444 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-940-572-2

Query Match   49.2%; Score 1119; DB 4; Length 444;
Best Local Similarity   49.8%; Pred. No. 7e-101;
Matches   213; Conservative   84; Mismatches 127; Indels   4; Gaps   4;

```
Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
        |||:||||||||| :  :||||||||||||||:|||||||||||:|||||:|| |||||
Db    5 YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYFTQRGLEQAGITILPFDEKNLDGDME 64

Qy   64 VIQGNAF-ASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHV 122
        :|||||| :  :|| || ||:::::||  || |:  :|  :| :||||||||||||:|||
Db   65 IIAGNAFRPDNNVEIAYADQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTGMLSHV 124

Qy  123 MNGDGKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI 182
        :: ||||||||||||||| |:::|||.| .::|:||:|| |||||:||||||||||||||:
Db  125 LSHITDTSFLIGDGTGRGSANAKYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTSL 184

Qy  183 NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFK-DSDDIYAQNIQITDKGT 241
        .|||.||:: | : :|:: |||:||.|||| ::|:||||||:  ::|::|| :  : |||
Db  185 EDVFNAFNDYAKQITKGLFVYGEDAELRKITSDAPIYYYGFEAEGNDFVASDLLRSTTGS 244

Qy  242 AFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNE 301
        :| |:: |: .:.|: : :| :.::|.:|| ||:|:: |.|::::.: | |||||||||:
Db  245 TFTVHFRGQNLGQFHIPTFGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTE 304

Qy  302 TTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSK 361
        .:: :: ||||:|||| ::|:: :||:|||...|:|||||||||:||  :|::|| :| :
Db  305 KIVNDTVIIDDFAHHPEIIATLDAARQKYPSSKEIVAVFQPHTFTRTIALLDDFAHALNQ 364

Qy  362 ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGD 419
        ||.|:| :|:||:||  : |. :::|:::|  .::.: :|.|:.::::.|||:::|||||
Db  365 ADAVYLAQIYGSAREVDHGDVKVEDLANKINKKHQVITVENVSPLLDHDNAVVVFMGAGD 424

Qy  420 IQKLQNAY 427
        ||.:: ::
Db  425 IQTYEYSF 432
```

RESULT 3
US-08-961-083-116
; Sequence 116, Application US/0896083
; Patent No. 6159469
; GENERAL INFORMATION:
;   APPLICANT: Choi et. al.
;   TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;   NUMBER OF SEQUENCES: 452
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,083
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 116:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 422 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-961-083-116

Query Match   45.8%; Score 1043; DB 3; Length 422;
Best Local Similarity   48.8%; Pred. No. 1.7e-93;
Matches   200; Conservative   80; Mismatches 126; Indels   4; Gaps   4;

```
Qy   22 HDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMVVIQGNAF-ASSHEIVRA 80
        |:|||||||||:||:|||||||||||||||||||||||||||||||||: :  :|||||
```

Left column — RESULT 4

```
Db    1 HQMGHKVQGSDVEKYYFTQRGLEQAGITILPFDEKNLDGDMEIIAGNAFRPDNNVEIAYA 60
Qy   81 HQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHVMNGDKKTSFLIGDGTGMG 140
Db   61 DQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTGMLSHVLSHITDTSFLIGDGTGRG 120
Qy  141 LPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDINDVFDAFQEMAHNVKKGI 200
Db  121 SANAKYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTSLEDVFNAFNDYAKQITKGL 180
Qy  201 IAWGDDEHLRKIEADVPIYYYGFK-DSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQ 259
Db  181 FVYGEDAELRKITSDAPIYYYGFEAEGNDFVASDLLRSITGSTFTVHFRGQNLGQFHIPT 240
Qy  260 YGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKQRRFNETTIANQVIVDDYAHHPRE 319
Db  241 FGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTEKIVNDTVIIDDFAHHPTE 300
Qy  320 ISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRE-NT 378
Db  301 IIATLDAARQKIPSKEIVAVFQPHTFTRTIALLDDFAHALNQADAVYLAQIYGSAREVDH 360
Qy  379 GALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDIQKLQNAY 427
Db  361 GDVKVEDLANKINKKHQVITVENVSPLLDHDNAVYVFMGAGDIQTYEYSF 410
```

RESULT 4
US-09-536-784-116
; Sequence 116, Application US/09536784
; Patent No. 6573082
  GENERAL INFORMATION:
     APPLICANT: Choi et. al.
     TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
     NUMBER OF SEQUENCES: 452
     CORRESPONDENCE ADDRESS:
        ADDRESSEE: Human Genome Sciences, Inc.
        STREET: 9410 Key West Avenue
        CITY: Rockville
        STATE: Maryland
        COUNTRY: USA
        ZIP: 20850
     COMPUTER READABLE FORM:
        MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
        COMPUTER: HP Vectra 486/33
        OPERATING SYSTEM: MSDOS version 6.2
        SOFTWARE: ASCII Text
     CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/536,784
        FILING DATE: 30-OCT-1997
        CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/961,083
        FILING DATE: OCT-30-1997
     ATTORNEY/AGENT INFORMATION:
        NAME: Michelle S. Marks
        REGISTRATION NUMBER: 41,971
        REFERENCE/DOCKET NUMBER: PB340P3
     TELECOMMUNICATION INFORMATION:
        TELEPHONE: (301) 309-8504
        TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 116:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 422 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
        MOLECULE TYPE: protein
  SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-09-536-784-116

Query Match   45.8%;   Score 1043;   DB 4;   Length 422;
Best Local Similarity 48.8%;   Pred. No. 1.7e-93;

Right column — RESULT 5

```
Matches  200;  Conservative  80;  Mismatches  126;  Indels  4;  Gaps  4;

Qy   22 HDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKBDMVVIQGNAF-ASSHEEIVRA 80
Db    1 HQMGHKVQGSDVEKYYFTQRGLEQAGITILPFDEKNLDGDMEIIAGNAFRPDNNVEIAYA 60
Qy   81 HQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHVMNGDKKTSFLIGDGTGMG 140
Db   61 DQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTGMLSHVLSHITDTSFLIGDGTGRG 120
Qy  141 LPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDINDVFDAFQEMAHNVKKGI 200
Db  121 SANAKYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTSLEDVFNAFNDYAKQITKGL 180
Qy  201 IAWGDDEHLRKIEADVPIYYYGFK-DSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQ 259
Db  181 FVYGEDAELRKITSDAPIYYYGFEAEGNDFVASDLLRSITGSTFTVHFRGQNLGQFHIPT 240
Qy  260 YGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKQRRFNETTIANQVIVDDYAHHPRE 319
Db  241 FGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTEKIVNDTVIIDDFAHHPTE 300
Qy  320 ISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRE-NT 378
Db  301 IIATLDAARQKIPSKEIVAVFQPHTFTRTIALLDDFAHALNQADAVYLAQIYGSAREVDH 360
Qy  379 GALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDIQKLQNAY 427
Db  361 GDVKVEDLANKINKKHQVITVENVSPLLDHDNAVYVFMGAGDIQTYEYSF 410
```

RESULT 5
US-09-107-532A-6914
; Sequence 6914, Application US/09107532A
; Patent No. 6583275
  GENERAL INFORMATION:
     APPLICANT: Lynn A Doucette-Stamm and David Bush
     TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
        ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
     NUMBER OF SEQUENCES: 7310
     CORRESPONDENCE ADDRESS:
        ADDRESSEE: GENOME THERAPEUTICS CORPORATION
        STREET: 100 Beaver Street
        CITY: Waltham
        STATE: Massachusetts
        COUNTRY: USA
        ZIP: 02354
     COMPUTER READABLE FORM:
        MEDIUM TYPE: CD/ROM ISO9660
        COMPUTER: PC
        OPERATING SYSTEM: <Unknown>
        SOFTWARE: ASCII
     CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/107,532A
        FILING DATE: 30-Jun-1998
     PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 60/085,598
        FILING DATE: 14 May 1998
        APPLICATION NUMBER: 60/051571
        FILING DATE: July 2, 1997
     ATTORNEY/AGENT INFORMATION:
        NAME: Ariniello, Pamela Deneke
        REGISTRATION NUMBER: 40,489
        REFERENCE/DOCKET NUMBER: GTC-012
     TELECOMMUNICATION INFORMATION:
        TELEPHONE: (781)893-5007
        TELEFAX: (781)893-8277
  INFORMATION FOR SEQ ID NO: 6914:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 291 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein

```
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;         ORGANISM: Enterococcus faecium
;       FEATURE:
;         NAME/KEY: misc feature
;         LOCATION: (B) LOCATION 1...291
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6914:
US-09-107-532A-6914

  Query Match             42.5%;  Score  967;  DB 4;  Length 291;
  Best Local Similarity   63.8%;  Pred. No. 2.7e-86;
  Matches  178;  Conservative  43;  Mismatches  58;  Indels  0;  Gaps  0;

Qy    4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKILPFDANNIKEDMV   63
         ||||||||||||||| : : :::    ||||| :  : : :  : ||||: |:||:|
Db   13  YHFVGIKGSGMSSLALVLHEQGLNVQGSDIEKYFFTQRDLEKANITILPFNADNVKPGMT   72

Qy   64  VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM  123
         :|:||||  |||||  ||:|  : |   || : |: ||||:|||| ||||||||||||||
Db   73  IIAGNAFPPDSHEEIQRAKELGLEVIRYHDFIGHFIQNYTSIAVTGSHGKTSTTGLLSHVL  132

Qy  124  NGDKKTSFLIGDGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN  183
          |    : |||||||||  | :|||:|:||||||||||:|||||||||||||||||||| :
Db  133  SGVRPTSYLIGDGTGHGDPQAEFFSFRACEYRRHFLAYSPDYAIMTNIDFDHPDYYTSID  192

Qy  184  DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF  243
         |||:||| ||  |||: ||||| |||:||||:|||||| :  :||||: :||:: :|:||
Db  193  DVFTAFQTMAGQVKKAIPAYGDDAYLRKLKANVPIYYYGVTENDDIQARNIERTTSGSAF  252

Qy  244  DVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDV  282
         ||| |:||| ||  | |:|||:|||| |||: :||||| :
Db  253  DVYHGDEFVGHFTVPAFGKHDIINALGVIAVAVEKLDL  291


RESULT 6
US-09-328-352-5023
;  Sequence 5023, Application US/09328352
;  Patent No. 6552958
;  GENERAL INFORMATION:
;    APPLICANT: Gary L. Breton et al.
;    TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;    TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;    FILE REFERENCE: GTC99-03PA
;    CURRENT APPLICATION NUMBER: US/09/328,352
;    CURRENT FILING DATE: 1999-06-04
;    NUMBER OF SEQ ID NOS: 8252
;  SEQ ID NO 5023
;    LENGTH: 492
;    TYPE: PRT
;    ORGANISM: Acinetobacter baumannii
US-09-328-352-5023

  Query Match             25.2%;  Score  574;  DB 4;  Length 492;
  Best Local Similarity   34.1%;  Pred. No. 1.7e-47;
  Matches  154;  Conservative  86;  Mismatches  171;  Indels  40;  Gaps  14;

Qy    3  HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKI-LPFDANNIKED   61
          |||||| |:| | ::: :  |||: ||||| ::  :     |:  |||  ||||  ||:
Db   34  HIHFVGIGGAGMCGIAEVLANQGYVTGSDIKASKTTQ-QLEENGIKVVIGHEAENIKNA   92

Qy   62  MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH  121
           :::  || : : |  :: :   :|::   : |: :  |: |  : |||||||||||||:
Db   93  NVLVVSTAIDPENPEVKAAIEQRIPIVRRAEMLGELMRYRHGIAVAGTHGKTTTSLLTT  152

Qy  122  VMNGDK-KTSFLIG-----DGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDH  175
          :  :|  :||:|       ||||  |||:|||:|:||||||||| : :||||||||| ||
Db  153  MLAEENLDPTYVIGGLLNSTGVNRAALGESRFIVARADESDASFLYLQPMAAIVTNIDADH  212

Qy  176  PD-YFKDINDVFDAFQEMAHNVK-KGI-IAWGDDEHLRKIEADV--PIYYYGFKDSDDIY  230
          |  :|:|::|    |: : ::   ||:  : |||:  : : |    |:::||:::::|:
Db  213  MDTYEGSFDKLKDTFVQFLHNLPFYGLAVVCGDDANIREILPRVGRPVITYGFNEDNDIR  272

Qy  231  AQNIQITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALE  290
         || || ||||  ||||||||| ||||||| ||||:|||: :::||||||: :  |:|::
Db  273  AIDVEQDGMRSHFTVLRKGREPLRLTINQPGLHNVLNALAAIGVATDEGVSDEAISRALK  332

Qy  291  TFGGVKRRFN-----ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTF  345
          : || :|||     | :    :  |||| || :: :|:  |:  :    :  ::|||: :
Db  333  GPSGVGRRFQVQGEFELGEGNVKLVDDYGHHPKEVEATIKAARQSHPDRRLVMLFQPHRY  392

Qy  346  SRTQAFLNEFAESLSKADRVFLCEIFGS------IRENTGALTIQ----DLIDKIE      391
         |||   :::| : ||: |      :: :      ||:||||:  :     ||||: |
Db  393  SRTRDCFDDFIEVLSQVDQLLLLEVYPAGEKPIVGADSRTLARSIRLRGQVEPILIDPVE  452

Qy  392  GASLINEDSI-NVLEQFDNAVILFMGAGDI                               420
         |      : |  ||:: : ::|:: : |::
Db  453  G---NLQNIMQNVLQ--PNDLLLTQGAGNV                               477


RESULT 7
US-09-540-236-2712
;  Sequence 2712, Application US/09540236
;  Patent No. 6673910
;  GENERAL INFORMATION:
;    APPLICANT: Gary L. Breton et al.
;    TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAI
;    TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;    FILE REFERENCE: 2709.2005-001
;    CURRENT APPLICATION NUMBER: US/09/540,236
;    CURRENT FILING DATE: 2000-04-04
;    NUMBER OF SEQ ID NOS: 3840
;  SEQ ID NO 2712
;    LENGTH: 493
;    TYPE: PRT
;    ORGANISM: M.catarrhalis
US-09-540-236-2712

  Query Match             23.4%;  Score  532;  DB 4;  Length 493;
  Best Local Similarity   32.1%;  Pred. No. 2.2e-43;
  Matches  146;  Conservative  78;  Mismatches  185;  Indels  185;  Gaps  13;

Qy    5  HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKI-LPFDANNIKEDMV   63
         ||||| |:| | ::  : | |: ||||| ::|:  :  |   |||   |:||: ||:|:||
Db   34  HFVGIGGAGMCGIAEVLANQGYVTGSDIKANAMTE-RLENLGVTVHVGHDASNIKMADV   92

Qy   64  VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM  123
          : : |: :: |  :| :  : ::  :  : | :: |: : | || |||||||||||:::
Db   93  VVVSSAIDRQNPEIRAALKAHIPVVRRADMLGELMRYRHGIAVAGAHGKTTTTSLLTMML  152

Qy  124  -NGDKKTSFLIG-----DGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDHPD  177
           |    | |::     || :  ||::|:: :::|||||||||| :  |||||||| :||
Db  153  TERAGLDPTYVIGGKLNASGKNAALGASRYIVAEADESDASFLSLRPMACVVTNIDEDHME  212

Qy  178  -YFKDINDVFDAFQEMAHNVK-KGI-IAWGDDEHLRKIEADV--PIYYYGFKDSDDIYAQ  232
          ::|::::    |: : :: :  :|  : |||::  ||: |   |:::||  |:|:::
Db  213  TYEGSFDKLKQAVVQFLHNMPFFYGLAVLCGDDKELYAMIDDIARPVITYGLEKHNDV---  269

Qy  233  NIQITDKGTAFDVYVDGEFYDHFLSPQY----------GDHTVLNALAVIAISYLEKL    280
         ::  : :          :   : |          |:|||:|||:: :|  ||:
Db  270  ------QAVDVIADGT-KTHFTVLRKDKKPLPITLNIPGIHNVYNALGAITMATDEGV  320

Qy  281  DVTNIKEALETFGGVKRRFNET------TIANQVIVDDYAHHPREISATIETARKKYPH  333
         |   | :| |: || :||||:       | |:   |||| ||::: :| |  |::  |
Db  321  SDKAICQAVEKFAGVGRRFENNGSYPLTDGSGDVVLIDDYGHHFTEIAMTIKAARQSYPD  380

Qy  334  KEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEGA  393
          : : :||||: |||  | :|| :||:: |  ||  ::   |:|| | ||||:||::  |
Db  381  RRLVMMFQPHRYSRTRDCFSEFVNVLSQVDKLILLDVYSAGEELKGATSNDLARSLRER  440

Qy  394  SLINEDSINV--LEQFDNAVILFMGAGDIQKLQNA                           426
          |:|: : ||  :: : : :::: : |:::|:|
Db  441  GQVEPIVLNVNDKEQITQVLKITILNANDLLMTQGA                          475


RESULT 8
US-09-489-039A-12280
```

```
; Sequence 12280, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12280
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12280

Query Match             21.9%;  Score 499;  DB 4;  Length 492;
Best Local Similarity   30.0%;  Pred. No. 3.6e-40;
Matches  137;  Conservative  86;  Mismatches 192;  Indels  42;  Gaps  13;

Qy    3 HHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKED 61
        | |||||| | : |||:|:  :||| | |:::  ||||| ::||
Db   22 HIHFVGIGGAGMCGIAEVLANEGYQISGSDLAPNPVTQ-QLSQLGATIYFNHRPENVRDA 80

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSH 121
        : :: ||| : :||| | : || ||||:  ||||||||||:: ||
Db   81 SVVVVSAISADNPEIVAAHEARIPVIRRAEMLAELMRFRHGIAIAGTHGKTTTAMVSS 140

Qy  122 V------MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNI 171
        :       : ||  : || |||||:
Db  141 IYAEAGLDPTFVNG---GLVKAAGVHARLGHSRYLIAEADESDASFLHLQPMVAIVTNI 196

Qy  172 DFDHPD-YFKDINDVFDAFQEMAHNVK--KGIIAWGDDEHLRKIEADV--PIYYYGFKDS 226
        : ::    :  : | : : || :|||||| | :|:
Db  197 EADHMDTYHGDFENLKQTFINFLHNLPFYGRAVMCVDDPVIRELLPRVGRQITTYGFSDD 256

Qy  227 DDIYAQNI-QITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNI 285
        || :  :: ::  |||:   ||||   ::|
Db  257 ADVRVEDYRQVGAQGHFRLVRQDKAILQVTLNAP-GRHNALNAAAAVAVATEEGIDDRAI 315

Qy  286 KEALETFGGVKRRFN---ETTIA-------NQVIVDDYAHHPREISATIETARKKYPHKE 335
        : || ||| :||| :||| :||| |:|: | |: :
Db  316 LRALESFQGTGRRFDFLGEFPLAEVNGKPGSAMLIDDYGHGHPTEVDATIKAARAGWPDKN 375

Qy  336 VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALFIQDLIDKIEGASL 395
        :| |||| : ||: | | |:::  | |:: ::| ||||:
Db  376 LVMVFQPHRYTRTRDLYDDFANVLTQVDALLMLDVYPAGEAPIPGADSRSLCRTIRGRGK 435

Qy  396 IN----EDSINVLEQF-----DNAVILFMGAGDIQKL 423
        ::     :::   ||  |||:|||||||
Db  436 VDPILVPDSAQAAEMLASVLTGNDLIVLVQGAGNIGKI 472
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24196

Query Match             21.5%;  Score 490;  DB 4;  Length 513;
Best Local Similarity   30.9%;  Pred. No. 2.9e-39;
Matches  147;  Conservative  77;  Mismatches 195;  Indels  56;  Gaps  14;

Qy    5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKEDMV 63
        |||||| | : |||:|:  :||| | |:::  :|||:| :|
Db   52 HFVGIGGAGMCGIAEVLLNLGYEVSGSDLKASAVTE-RLEKFGAQIFIGHQAENADGADV 110

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHVM 123
        :: :: ||| : :||| | : || ||||:  ||||||||||::: : :
Db  111 LVVSSAINRANPEVASALERRIPVVPRAEMLAELMRYRHGIAVAGTHGKTTTSLIIASVF 170

Qy  124 -NGDKKTSFLIGD-----GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH-P 176
          : ||  : || |||||:  ::| :||:| : ||| ::| || :|:: :
Db  171 AAGGLDPTFVIGGRLNAAGTNAQLGASRYLVAEADBSDASFLHLQPMVAVVNIDADHMA 230

Qy  177 DYFKDINDVFDAFQEMAHNVK-KGI-IAWGDDEHLRKIEADV--PIYYYGFKDSDDIYAQ 232
        ||:  : | : : || :|||||:  | :|:  ||| ::|
Db  231 TYGGDFNKLKKTFVEFLHNLPFYGLAVMCVDDPVVREILPQIARPTVTYGLSEDADVRAI 290

Qy  233 NIQITDKGTAFDVYVDGEFYDHFLSPQY--------GDHTVLNALAVIAISYLEKLDV 282
        || : :  : | | :: ::|||:|:
Db  291 NIRQEGMRTWFTV----------LRPEREPLDVSVNMPGLHNVLNSLATIVIATDEGISD 340

Qy  283 TNIKEALETFGGVKRRFN-----ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVV 337
        : : || ||| :|||| :||:| :|:||||: :: | |:|
Db  341 EAIVQGLSGFQGVGRRFQVYGELQVEGGSVMLVDDYGHHPREVAAVIKAIRGGWPERRLV 400

Qy  338 AVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLIN 397
        :|||| : ||: | | | :: | |:: :| | |||:|||||:  | | :|:
Db  401 MVYQPHRYTRTRDLYEDFVQVLGEANVLLMEVVPAGEEPIPGADSRQLCHSIRQRGQL- 459

Qy  398 EDSINVLBQFDNA------------VILFMGAGDI---QKLQNAYLDKLGMKNA 436
        :::   ||  |||:|||||:    :| ::: || :|:
Db  460 -DPIYFERDADLAPLVKPLLRAGDILLCQGAGDVGGGLAPQLIKNPLFAGKGGKGA 513

RESULT 10
US-09-543-681A-6951
; Sequence 6951, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6951
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6951

Query Match             20.9%;  Score 475;  DB 4;  Length 488;
Best Local Similarity   28.0%;  Pred. No. 8e-38;
Matches  132;  Conservative  86;  Mismatches 183;  Indels  70;  Gaps  11;

Qy    3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDM 62
         | |||||| | : |||:|:  :||| | |:::  :|||:|
Db   22 HIHFVGIGGAGMGGIAEVLANEGYAISGSDLAPNVVTQQLVALGATIYFNHRPENIRDAS 81

Qy   63 VVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHV 122
        ||: :: ||| : :||| | : || ||||:  ||||||||||::: :
Db   82 VVVVSTAISADNPEIIAAREARIPVIRRAEMLAELMRYRHGVAIAGTHGKTTTAMISNI 141
```

RESULT 9
US-09-252-991A-24196
```
; Sequence 24196, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24196
; LENGTH: 513
```

```
Qy  123 --------MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNID 172
              :|:  ||:|:||           ||    |::|||:
Db  142 YAQAGLDPTFVNG---GLVKSAGTHARLGCSRYLIAEADESDASFLHLQPMVAVVTNIE 197

Qy  173 FDHPD-YFKDINDVFDAFQEMAHNVKKGIIAWG------DDEHLRKIEADVPIYY--YGF 223
              |: ||::: : :||                     ||
Db  198 ADHMDTYHGNFDNLKETFITFLHNLP----FYGRAVMCIDDDVIRSIIPKVGRYITTYGF 253

Qy  224 KDSDDIYAQNIQITDKGTA--------FDVYVDGEFYDHFLSPQYGDHTVLNALAV 271
              |::
Db  254 SEDADVRITHYE--QKGAQGFFTISREDMPDIDVVLNAP----------GRHNALNATAA 301

Qy  272 IAISYLEKLDVTNIKEALETFGGVKRRFN------ETTIANQVIVDDYAHHPREIS 321
              |:
Db  302 VAVATEGIADEHILAALLNFQGTGRRFDFLGNFSLEHVNGQEGEVMLVDDYGHHPTEVD 361

Qy  322 ATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGS-------- 373
              ||||
Db  362 ATIKAARAGWPDKRLVMLFQPHRYTRTRDLYEDFATVLNQVDILLLTDVYAAGEAPIPGA 421

Qy  374 -IRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDIQKL 423
              |:: :::
Db  422 DSRSLCRTIRQRGKLDPIWVSDVENISSILAGVLTDNDLVLVQGAGNIGKI 472
```

RESULT 11
US-09-198-452A-978
; Sequence 978, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;   TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;   TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
;   TITLE OF INVENTION: and treatment of infection
;   FILE REFERENCE: 9710-003-999
;   CURRENT APPLICATION NUMBER: US/09/198,452A
;   CURRENT FILING DATE: 1998-11-24
;   NUMBER OF SEQ ID NOS: 6849
;   SEQ ID NO 978
;   LENGTH: 812
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-198-452A-978

```
Query Match     20.7%;  Score 470;  DB 4;  Length 812;
Best Local Similarity   28.2%;  Pred. No. 5.5e-37;
Matches 131;  Conservative 89;  Mismatches 181;  Indels 64;  Gaps 16;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDI-ENYVFTEVALRNKGIKILP-FDANNIKED 61
        ||| | |:| ||:|   |  | ::: ||||
Db   10 YHFGIGGIGMSALAHILLDRGYEYSGSDLYESY--TIESLKAKGARCFSSGHDSSHVPHD 67

Qy   62 MVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
        :|: || :|   ||:::
Db   68 AVVYSSSIAPDNVEYLTAIQRSRLLHRAELLSQLMEGYESILVSGSHGKTGTSSLIRA 127

Qy  122 V-MNGDKKTSFLIGDGTGMGLP-------ESDYFAFRACEYRRHFLSYKPDYAIMTNI 171
        : :|
Db  128 IFQEAQKDPSYAIG-----GLAANCLNGYSGSSKIFVAEADESDGSLKHYTPRAVVITNI 182

Qy  172 DFDH-PDYFKDINDVFDAFQEMAHNVKKGIIAWGDDEHLRKI--EADVPI-------YYY 221
        | :|
Db  183 DNEHLNNYAGNLDNLVQIQDFSRKV---------TDLNKVFYNGDCPILKGNVQGISY 232

Qy  222 GFKDSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQY--------GDHTVLNALAVIA 273
        |: ||
Db  233 GY--SPECQLHIVSYNQKAW-----QSHFSFTFLGQEYQDIELNLPGQHNAANAAAACG 284

Qy  274 ISYLEKLDVTNIKEALETFGGVKRRFNETTIANQ-VIVDDYAHHPREISATIETARKKYP 332
        ::  | || ||| :||: |||
Db  285 VALTFGIDINIIRKALKKFSGVHRRLERKNISESFLFLEDYAHHPVEVAHTLRSVRDAVG 344

Qy  333 HKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEG 392
```

```
       ::|:||||| |||| ||| ||
Qy  393 ASLI-----NEDSINVLEQF--DNAVILFMGAGDIQKLQNAYLD 429
        :|  :   ::  :  : ||:||:  |||:||
Db  405 SSYVHCCVPHGDIVDYLRNYIRIHDVCVSLGAGNIYTIGEALKD 449
```

RESULT 12
US-09-328-352-4942
; Sequence 4942, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;   APPLICANT: Gary L. Breton et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;   TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: GTC99-03PA
;   CURRENT APPLICATION NUMBER: US/09/328,352
;   CURRENT FILING DATE: 1999-06-04
;   NUMBER OF SEQ ID NOS: 8252
;   SEQ ID NO 4942
;   LENGTH: 475
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-4942

```
Query Match     18.3%;  Score 417;  DB 4;  Length 475;
Best Local Similarity   29.2%;  Pred. No. 3.6e-32;
Matches 135;  Conservative 75;  Mismatches 202;  Indels 50;  Gaps 15;

Qy    3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNIK-- 59
        | |::| ||:|:| ||: ||| |||| ||::  :||
Db   20 HLHILGICGTFMGSLALLARDLGHKVTGSDSNVYPPMSTQLENAGIELMQGYDRSHLQPH 79

Qy   60 EDMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSV-AVTGAHGKTSTTGL 118
        ::| :: ||  :  ||||  ::| ||::
Db   80 PDLVIV-GNAMKRGIDAVEYMLNEGLPYISGPQFLADHVLQGKHVLGVAGTHGKTTTTM 138

Qy  119 LSHVMN-GDKKTSFLIGDGTGMGLPES------DYFAFEACEY------RRHFLSYKPD 164
        |:: |||:
Db  139 LAWVLDQAGLNPGFLIG-GVPLGFSESARLGGGKYFVVEADEYDSAFFDKRSKFVHYHPK 197

Qy  165 YAIMTNIDFDHPDYFKDINDVFDAFQEMAHNV--KKGIIAWGDDEHLRKI--EADVPIY 219
        ||: |||::::||
Db  198 TAILNNLEFDHADIFDDLAAIQKQFHHLVRTIPSEGRIIAPITETHIDEVLEMGCWTPVI 257

Qy  220 YYGFKDSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEK 279
        :|:||  :||| |  |:|:||||:||:|
Db  258 RTSLEANEKAALSAELISIDGSHFKVLENGNVIGEVKWSWTGQHSVANALATIAAAQHVG 317

Qy  280 LDVTNIKEALETFGGVKRPNETTIANQV-IVDDYAHHPREISATIETARKKYPHKEVVA 338
        |:| | || :||| ::|
Db  318 VSLEKACEALSNFGGVKRRMELLGTINGIEVYDDFAHHPTAIDTTLDGARKRLGERRIWA 377

Qy  339 VFQPHTFS-RTQAFLNEFAESLSKADRVFLCE-----------IPGSIRENTGALTIQDL 386
        :|::||  ::|| ||||||||||||:|
Db  378 IIEPRSNTMRMGSHKDGLAHSARLADEVIWYQPEGLDWDLQPVIEAATNHAQVSRSLDEI 437

Qy  387 IDKIEGASLINEDSINVLEQFDNAVILFMG--AGDIQKLQNA 426
        ||:
Db  438 IDRI-------------VNEAGEGDAVVIMSNGGFGGLHQKLMSA 469
```

RESULT 13
US-09-540-236-2891
; Sequence 2891, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;   APPLICANT: Gary L. Breton et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
;   TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 2709.2005-001
;   CURRENT APPLICATION NUMBER: US/09/540,236

```
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2891
; LENGTH: 473
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2891

Query Match        17.0%;    Score 386.5;  DB 4;   Length 473;
Best Local Similarity   27.8%;    Pred. No. 3.5e-29;
Matches  136;  Conservative   82;  Mismatches 177;  Indels  95;  Gaps  22;

Qy    3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNI--K  59
          |:||:|||:||||::|||||||||||||||:||||||:::|| ::
Db   15 HIHILGICGTFMGSLALLARDLGHTVTGSDANIYPPMSTQLADAGVEVWEGVHAEHLIPT  74

Qy   60 EDMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQII--DQYTSVAVTGAHGKTSTTG 117
Db   75 PDLIVV-GNACKRGMPAIEYMLNERLPYTSGPQFLYETVLKDRHV-LAVAGTHGKTTTTT 132

Qy  118 LLSHVMN-GDKKTSFLIGDGTGMGLPESD-----------YPAFACEY----- 154
Db  133 MLAWILQFSGIDTGFLIG---GVPLVNTDDRRLSLAFQHSSYLGKQFFVIEADEYDSAFF 189

Qy  155 --RRHFLSYKPDYAIMTNIDFDHPDYFKDINDVFDAFQEMAHNV-KKGIIAWGDDEHLRK 211
Db  190 DKRSKFVHYRPTTAILNNLEYDHADIFADLDAIQTQFHHMIRMIPSKGQII--------- 240

Qy  212 IEADVPIYYYGFKDSDDIYAQNIQIT-DKG---TAFDVYVDGE-------FYD 253
Db  241 IPPANTP------------------------SLETTLDKGVWTTVVRTSLNGDAEWQAKLDANDGSSFWV 285

Qy  254 HFLS-------POYGDHTVLNALAVIAISYLEKLDVTNIKEALETPGGVKRRFNETTIAN 306
Db  286 HHNNTSAMVRWSMSGLHNVNNALTAIAAAHHVGVSVEMACTALSHFGGIKRRMELIGDVD 345

Qy  307 QVIV-DDYAHHPREISAITETARKKYPHKEVVAVFQPHTFS-RTQAFLNEFAESLSKADR 364
Db  346 DIIVFDDFAHHPTAISTTLDGAKKRLTPRRIWAIIEPRSNTMKLGSHRPHLAASAAIADQ 405

Qy  365 VFLCEIFG---SIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDIQ 421
Db  406 VIWYEPQGLTWGLKEAIGSTPNQRVLDSIN--AII--EHIKTYAKAGDAIII-MSNGDFE 460

Qy  422 KLQNAYLDKL 431
Db  461 NIHGRLLDAL 470

RESULT 14
US-09-543-681A-5705
; Sequence 5705, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5705
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5705

Query Match        16.3%;    Score 371;  DB 4;   Length 471;
Best Local Similarity   29.8%;    Pred. No. 1.1e-27;
Matches  120;  Conservative   66;  Mismatches 161;  Indels  56;  Gaps  17;
```

```
Qy    3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNIK-- 59
          |:||:|||:||||::|||||||||||||||:||||||:::|| ::
Db   14 HIHILGICGTFMGSLAILARAKGHKVTGSDANVYPPMSTLLENQGIELIEGYDPQQLEPA 73

Qy   60 EDMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFL-GQIIDQYTSVAVTGAHGKTSTTGL 118
Db   74 PDMVII-GNAMTRGNPCVEAVLEKGLPYTSGPQWLHDYILPERWVLAVAGTHGKTTTAGM 132

Qy  119 LSHVMNG-DKKTSFLIGDGTG-----MGLPESDYAFAEACEY------RRHFLSYKPDY 165
Db  133 LAWILEDCGYQPGFLIGGVPGNFQVSAQLGESPFFVIEADEYDSAFFDKRSKFVHYSPRT 192

Qy  166 AIMTNIDFDHPDYFKDINDVFDAFQEMAH--------------NVKK--GIIAWGDDEH 208
Db  193 LILNNLEFDHADIFPDDLTAIQKQFIHHLVRIVPGSGKIIMPDNDLNLKQTIGMGCWSEEE- 251

Qy  209 LRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNA 268
Db  252 ------------YTG--ETGDWQAK--KISNDSSHFAVFHKGEQVGEVCWGLSGEHNMQNG 296

Qy  269 L-AVIAISYLEKLDVTNIKEALETPGGVKRRFNETTIANQY-IVDDYAHHPREISAIIET 326
Db  297 LMAIVAAHHVGVLPV-DACAALNKFINARRLELRGEVNQVSVYDDFAHHPTAILATLEA 355

Qy  327 ARKKY-PHKEVVAVFQPHTFSRTQAP-LNEFAESLSKADRVFL 367
Db  356 LRSKVGGTARIIAVLEPRSNTMKMGISKDDIAPALGRADEVFL 398

RESULT 15
US-09-489-039A-8987
; Sequence 8987, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8987
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8987

Query Match        15.8%;    Score 358.5;  DB 4;   Length 458;
Best Local Similarity   27.2%;    Pred. No. 1.8e-26;
Matches  129;  Conservative   78;  Mismatches 202;  Indels  65;  Gaps  18;

Qy    5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNI--KED 61
Db    5 HILGICGTFMGGLAMLARSLGHEVTGSDANVYPPMSTLLENQGIDLIQGYDPSQLEPRPD 64

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQII--DQYTSVAVTGAHGKTSTTGLL 119
Db   65 LVII-GNAMTRGNPCVEAVLENNIPYMSGPQWLHDFVLRDRWV-LAVAGTHGKTTTAGMA 122

Qy  120 SHVMNG-DKKTSFLIGDGTG-----MGLPESDYAFRACEY------RRHFLSYKPDYA 166
Db  123 TWILEACGYKPGFVIGGVPGNFDVSARLGDSPFFVIRADEYDCAFFDKRSKFVHYCPRTL 182

Qy  167 IMTNIDFDHPDYFKDINDVFDAFQEMAHNV--KKGIIAWGDDEHLRKIEADVPIYYYGFK 224
Db  183 ILNNLEFDHADIFPDDLKAIQKQFHHLVRIVPGGKIILPENDINLKQVMA------MGCW 236

Qy  225 DSDDI-----YAQNIQITDKGTAFDVVVDGEFYDHFLSPQYGDHTVLNLAVAIISYLEK 279
Db  237 SEQELVGEQGHWQAKKLNADASQWEVLLDGEKVGEVKWALVGEHNMHNGLMAIAAARHVG 296
```

```
Qy    280  LDVTNIKEALETFGGVKRRFNETTIANQVIV-DDYAHHPREISATIETARKKY-PHKEVV 337
                :    .   || : | . :||    .||| ||||| | .||:: |         ::
Db    297  VLPADAANALGSFINARRRLELRGEANGVTVIDDFAHHPTAILATLQALRGKVGGTARIL 356

Qy    338  AVFQPHTFSRTQAFL-NEFAESLSKADRVFLCE---IFGSIRENTGAL----------T 382
             || : |:  :     .:   ||||| : | .     |   |        |
Db    357  AVLEPRSNTMKMGVCKDDLAPSLGRADEVFLLQPQHIPWQVAEVAEACVQPAHWSADVDT 416

Qy    383  IQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDIQKLQNAYLDKLGMKNA 436
             : : ::         :  : . | : || |      ||||  |||
Db    417  LAEMV-----------VKTAHPGDH--ILVMSNGGFGGIHQKLLDKLLAQKAA 455
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      June  3, 2004, 14:40:41 ; Search time 48 Seconds
                                (without alignments)
                                2561.354 Million cell updates/sec

Title:       US-10-712-713-2
Perfect score:  2275
Sequence:    1 MTHYHFVGIKGSGMSSLAQI..........GDIQKLQNAYLDKLGMKNAF 437

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:    1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters:    1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   Published Applications AA:*
             1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
             2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
             3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
             4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
             5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
             6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
             7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
             8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
             9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
            10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
            11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
            12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
            13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
            14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
            15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
            16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
            17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
            18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                     %
Result         Query
No.    Score   Match  Length DB  ID                       Description
-----  -----   -----  ------ --  ----------------------   -----------
  1    2275    100.0   437    9  US-09-925-637-2          Sequence 2, Appli
  2    2275    100.0   437   14  US-10-084-205-2          Sequence 2, Appli
  3    2275    100.0   437   16  US-10-712-713-2          Sequence 2, Appli
  4    2274    100.0   437   12  US-10-282-122A-43780     Sequence 43780, A
  5    2267     99.6   437    9  US-09-815-242-5297       Sequence 5297, Ap
  6    2267     99.6   444    9  US-09-815-242-12293      Sequence 12293, A
  7    1985     87.3   437   12  US-10-282-122A-71178     Sequence 71178, A
  8    1470     64.6   436   12  US-10-282-122A-46331     Sequence 46331, A
  9    1438.5   63.2   445   12  US-10-282-122A-57306     Sequence 57306, A
 10    1433.5   63.0   456    9  US-09-815-242-10870      Sequence 10870, A
 11    1396     61.4   444   12  US-10-282-122A-57651     Sequence 57651, A
 12    1342     59.0   447   12  US-10-282-122A-60988     Sequence 60988, A
 13    1139     50.1   442   12  US-10-282-122A-74395     Sequence 74395, A
 14    1132.5   49.8   443   12  US-10-282-122A-72179     Sequence 72179, A
 15    1119     49.2   444    9  US-09-815-242-13507      Sequence 13507, A
 16    1119     49.2   444   12  US-10-282-122A-74073     Sequence 74073, A
 17    1064     46.8   246   12  US-10-282-122A-71315     Sequence 71315, A
 18    1043     45.8   422    9  US-09-765-272-116        Sequence 116, App
 19     690.5   30.4   461   12  US-10-282-122A-52628     Sequence 52628, A
 20     641.5   28.2   458   12  US-10-282-122A-51682     Sequence 51682, A
 21     585.5   25.7   450   12  US-10-282-122A-53624     Sequence 53624, A
 22     574     25.2   482   12  US-10-282-122A-44665     Sequence 44665, A
 23     559     24.6   469   12  US-10-282-122A-61427     Sequence 61427, A
 24     548     24.1   506   12  US-10-282-122A-66060     Sequence 66060, A
 25     545     24.0   469   12  US-10-282-122A-65124     Sequence 65124, A
 26     538     23.6   468   12  US-10-282-122A-47281     Sequence 47281, A
 27     532     23.4   486   12  US-10-282-122A-63073     Sequence 63073, A
 28     509     22.4   465   12  US-10-282-122A-50582     Sequence 50582, A
 29     507     22.3   491   12  US-10-282-122A-55678     Sequence 55678, A
 30     506.5   22.3   468   12  US-10-282-122A-51373     Sequence 51373, A
 31     503.5   22.1   486    9  US-09-738-626-5862       Sequence 5862, Ap
 32     502     22.1   479   12  US-10-282-122A-69707     Sequence 69707, A
 33     499     21.9   491   12  US-10-282-122A-59579     Sequence 59579, A
 34     496     21.8   833   12  US-10-282-122A-55085     Sequence 55085, A
 35     493.5   21.7   477   12  US-10-282-122A-54019     Sequence 54019, A
 36     491     21.6   491    9  US-09-741-669-455        Sequence 455, App
 37     491     21.6   491    9  US-09-815-242-10032      Sequence 10032, A
 38     491     21.6   491   12  US-10-282-122A-56416     Sequence 56416, A
 39     490.5   21.6   491   12  US-10-282-122A-76022     Sequence 76022, A
 40     490     21.5   480    9  US-09-815-242-12008      Sequence 12008, A
 41     490     21.5   480   12  US-10-282-122A-66650     Sequence 66650, A
 42     488.5   21.5   482   12  US-10-282-122A-68321     Sequence 68321, A
 43     488     21.5   476   12  US-10-282-122A-73039     Sequence 73039, A
 44     487     21.4   481   12  US-10-282-122A-76478     Sequence 76478, A
 45     484.5   21.3   491    9  US-09-815-242-14079      Sequence 14079, A
```

ALIGNMENTS

```
RESULT 1
US-09-925-637-2
; Sequence 2, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
;  APPLICANT: Choi
;  TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
;  FILE REFERENCE: PB560
;  CURRENT APPLICATION NUMBER: US/09/925,637
;  CURRENT FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/23773
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US 60/151,933
;  PRIOR FILING DATE: 1999-09-01
;  PRIOR APPLICATION NUMBER: US 08/781,986
;  PRIOR FILING DATE: 1997-01-03
;  PRIOR APPLICATION NUMBER: US 08/956,171
;  PRIOR FILING DATE: 1997-10-20
;  PRIOR APPLICATION NUMBER: US 60/009,861
;  PRIOR FILING DATE: 1996-01-06
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;  LENGTH: 437
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-925-637-2

Query Match             100.0%;  Score 2275; DB 9; Length 437;
Best Local Similarity   100.0%;  Pred. No. 2.9e-198;
Matches  437;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy       61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGARGKTSTGLLS 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437


RESULT 2
US-10-084-205-2
; Sequence 2, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
;   APPLICANT: Choi, Gil
;   TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
;   FILE REFERENCE: PB515P1
;   CURRENT APPLICATION NUMBER: US/10/084,205
;   CURRENT FILING DATE: 2002-02-28
;   PRIOR APPLICATION NUMBER: PCT/US00/23773
;   PRIOR FILING DATE: 2000-08-31
;   PRIOR APPLICATION NUMBER: 60/151,933
;   PRIOR FILING DATE: 1999-09-01
;   NUMBER OF SEQ ID NOS: 74
;   SOFTWARE: PatentIn Ver. 3.1
;   SEQ ID NO 2
;     LENGTH: 437
;     TYPE: PRT
;     ORGANISM: Staphylococcus aureus
US-10-084-205-2

  Query Match             100.0%;  Score 2275;  DB 14;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-198;
  Matches  437;  Conservative     0;  Mismatches     0;  Indels     0;  Gaps     0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
```

```
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437


RESULT 3
US-10-712-713-2
; Sequence 2, Application US/10712713
; Publication No. US20040082002A1
; GENERAL INFORMATION:
;   APPLICANT: Choi, Gil
;   TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
;   FILE REFERENCE: PB515P1
;   CURRENT APPLICATION NUMBER: US/10/712,713
;   CURRENT FILING DATE: 2003-11-14
;   PRIOR APPLICATION NUMBER: US/10/084,205
;   PRIOR FILING DATE: 2002-02-28
;   PRIOR APPLICATION NUMBER: PCT/US00/23773
;   PRIOR FILING DATE: 2000-08-31
;   PRIOR APPLICATION NUMBER: 60/151,933
;   PRIOR FILING DATE: 1999-09-01
;   NUMBER OF SEQ ID NOS: 74
;   SOFTWARE: PatentIn Ver. 3.1
;   SEQ ID NO 2
;     LENGTH: 437
;     TYPE: PRT
;     ORGANISM: Staphylococcus aureus
US-10-712-713-2

  Query Match             100.0%;  Score 2275;  DB 16;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-198;
  Matches  437;  Conservative     0;  Mismatches     0;  Indels     0;  Gaps     0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437
```

RESULT 4
US-10-282-122A-43780
; Sequence 43780, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Liangsu
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Malone, Cheryl
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Xu, H.
;   APPLICANT: Forsyth, R.
;   TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;   FILE REFERENCE: ELITRA.034A
;   CURRENT APPLICATION NUMBER: US/10/282,122A
;   CURRENT FILING DATE: 2003-02-20
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/230,335
;   PRIOR FILING DATE: 2000-09-06
;   PRIOR APPLICATION NUMBER: 60/230,347
;   PRIOR FILING DATE: 2000-09-09
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/267,636
;   PRIOR FILING DATE: 2001-02-09
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;   NUMBER OF SEQ ID NOS: 78614
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43780
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-282-122A-43780

Query Match   100.0%; Score 2274; DB 12; Length 437;
Best Local Similarity   99.8%; Pred. No. 3.6e-198;
Matches 436; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy   61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy   181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy   241 TAFDVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TAFDVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy   301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy   361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420

Qy   421 QKLQNAYLDKLGMKNAF 437
              |||||||||||||||||
Db   421 QKLQNAYLDKLGMKNAF 437

RESULT 5
US-09-815-242-5297
; Sequence 5297, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Kari L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Carr, Grant J.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
;   TITLE OF INVENTION: Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09/815,242
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5297
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5297

Query Match   99.6%; Score 2267; DB 9; Length 437;
Best Local Similarity   99.5%; Pred. No. 1.6e-197;
Matches 435; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFGANNIKE 60

Qy   61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy   181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

```
Db  128 HVMNGDKKTSFLIGDGTGMGLPESDYFAFFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 187
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 247
Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 307
Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 367
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLLNEDSINVLEQFDNAVLFMGAGDI 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||: |||||||
Db  368 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLLNEDSINVLEQFDNAVVLFMGAGDI 427
Qy  421 QKLQNAYLDKLGMKNAF 437
            |||||||||||||||||
Db  428 QKLQNAYLDKLGMKNAF 444
```

RESULT 7
US-10-282-122A-71178
; Sequence 71178, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; ; APPLICANT: Wang, Liangsu
; ; APPLICANT: Zamudio, Carlos
; ; APPLICANT: Malone, Cheryl
; ; APPLICANT: Haselbeck, Robert
; ; APPLICANT: Ohlsen, Kari
; ; APPLICANT: Zyskind, Judith
; ; APPLICANT: Wall, Daniel
; ; APPLICANT: Trawick, John
; ; APPLICANT: Carr, Grant
; ; APPLICANT: Yamamoto, Robert
; ; APPLICANT: Forsyth, R.
; ; APPLICANT: Xu, H.
; ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; ; FILE REFERENCE: ELITRA.034A
; ; CURRENT APPLICATION NUMBER: US/10/282,122A
; ; CURRENT FILING DATE: 2003-02-20
; ; PRIOR APPLICATION NUMBER: 60/191,078
; ; PRIOR FILING DATE: 2000-03-21
; ; PRIOR APPLICATION NUMBER: 60/206,848
; ; PRIOR FILING DATE: 2000-05-23
; ; PRIOR APPLICATION NUMBER: 60/207,727
; ; PRIOR FILING DATE: 2000-05-26
; ; PRIOR APPLICATION NUMBER: 60/230,335
; ; PRIOR FILING DATE: 2000-09-06
; ; PRIOR APPLICATION NUMBER: 60/230,347
; ; PRIOR FILING DATE: 2000-09-09
; ; PRIOR APPLICATION NUMBER: 60/242,578
; ; PRIOR FILING DATE: 2000-10-23
; ; PRIOR APPLICATION NUMBER: 60/253,625
; ; PRIOR FILING DATE: 2000-11-27
; ; PRIOR APPLICATION NUMBER: 60/257,931
; ; PRIOR FILING DATE: 2000-12-22
; ; PRIOR APPLICATION NUMBER: 60/269,308
; ; PRIOR FILING DATE: 2001-02-16
; ; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71178
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-71178

```
Db  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLLNEDSINVLEQFDNAVLFMGAGDI 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLLNEDSINVLEQFDNAVLFMGAGDI 420
Qy  421 QKLQNAYLDKLGMKNAF 437
            |||||||||||||||||
Db  421 QKLQNAYLDKLGMKNAF 437
```

RESULT 6
US-09-815-242-12293
; Sequence 12293, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; ; APPLICANT: Haselbeck, Robert
; ; APPLICANT: Ohlsen, Kari L.
; ; APPLICANT: Zyskind, Judith W.
; ; APPLICANT: Wall, Daniel
; ; APPLICANT: Trawick, John D.
; ; APPLICANT: Carr, Grant J.
; ; APPLICANT: Yamamoto, Robert T.
; ; APPLICANT: Xu, H. Howard
; ; TITLE OF INVENTION: Identification of Essential Genes in
; ; Prokaryotes
; ; TITLE OF INVENTION: Prokaryotes
; ; FILE REFERENCE: ELITRA.011A
; ; CURRENT APPLICATION NUMBER: US/09/815,242
; ; CURRENT FILING DATE: 2001-03-21
; ; PRIOR APPLICATION NUMBER: 60/191,078
; ; PRIOR FILING DATE: 2000-03-21
; ; PRIOR APPLICATION NUMBER: 60/206,848
; ; PRIOR FILING DATE: 2000-05-23
; ; PRIOR APPLICATION NUMBER: 60/207,727
; ; PRIOR FILING DATE: 2000-05-26
; ; PRIOR APPLICATION NUMBER: 60/242,578
; ; PRIOR FILING DATE: 2000-10-23
; ; PRIOR APPLICATION NUMBER: 60/253,625
; ; PRIOR FILING DATE: 2000-11-27
; ; PRIOR APPLICATION NUMBER: 60/257,931
; ; PRIOR FILING DATE: 2000-12-22
; ; PRIOR APPLICATION NUMBER: 60/269,308
; ; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12293
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12293

Query Match   99.6%; Score 2267; DB 9; Length 444;
Best Local Similarity 99.5%; Pred. No. 1.6e-197;
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    8 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFGANNIKE 67
Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 127
Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFFACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Query Match          87.3%;  Score 1985;  DB 12;  Length 437;
Best Local Similarity 84.7%;  Pred. No. 7.3e-172;
Matches  370;  Conservative  43;  Mismatches  24;  Indels  0;  Gaps  0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPPDANNIKE 60
        ||||||||||||||||||||||||||||||||||: :||||||||||||||||||||| :
Db    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIESVFTEVALRNKGIKILPPDANNITK 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        :|||||||||  :|||||||||:|||:: :|:||| :||||||||||||||||||||||
Db   61 EMVVIQGNAFPDNHEEIVRAHELKLDIIKYHDFLGHVINQYTSVAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||||||||||||||||||||| ||||||||||||||||||:||||||||||||||||||
Db  121 HVMNGDKKTSFLIGDGTGMGLPGSDYFAFEACEYRRHFLSYHPDYAIMTNIDFDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        :|:||:||||::||:|||||||||||||:|||:| ::|:||||:: |||||| ||||::|
Db  181 NIDDVYDAFQHMALNVKKGIIAWGDDEYLRKLDVDIPVYYYGFKETDDIYAKNIQITEKG 240

Qy  241 TAFDYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        |.||.|  |||||:||||||:| ||||||||||||||:: :.|:||| |:||||||||||
Db  241 TQFDVVIKGEFYDQFLSPQYGNHNILNALAVIAISYLEDMNVENIKEALITFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        ||  :.||||||||||||||||||||||||||:|:|||||||||||||||||||||||||
Db  301 ETKVSNQVIVDDYAHHPREISATIETARKKYPQKDVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||:|||||||||||||||.|.| ||:|.|: ||:|:|:||.|||:|||||||||||||||
Db  361 KADQVFLCEIFGSIRENTGDLITIEDLINRIDGSTLIDENSIDVLEKFDNAVILFMGAGDI 420

Qy  421 QKLQNAYLDKLGMKNAF 437
        |||:.||::||| ||:|
Db  421 QKLLKAYFEKLGVKNDF 437
```

RESULT 8
US-10-282-122A-46331
; Sequence 46331, Application US/10282122A
; Publication No. US2004002129A1
GENERAL INFORMATION:
;   APPLICANT: Wang, Liangsu
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Malone, Cheryl
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Kari
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Forsyth, R.
;   APPLICANT: Xu, H.
;   TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;   FILE REFERENCE: ELITRA.034A
;   CURRENT APPLICATION NUMBER: US/10/282,122A
;   CURRENT FILING DATE: 2003-02-20
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/230,335
;   PRIOR FILING DATE: 2000-09-06
;   PRIOR APPLICATION NUMBER: 60/230,347
;   PRIOR FILING DATE: 2000-09-09
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931

;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/267,636
;   PRIOR FILING DATE: 2001-02-09
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;   NUMBER OF SEQ ID NOS: 78614
;   SOFTWARE: PatentIn version 3.1
SEQ ID NO 46331
;   LENGTH: 436
;   TYPE: PRT
;   ORGANISM: Bacillus anthracis
US-10-282-122A-46331

```
Query Match          64.6%;  Score 1470;  DB 12;  Length 436;
Best Local Similarity 64.7%;  Pred. No. 5.4e-125;
Matches  279;  Conservative  59;  Mismatches  93;  Indels  0;  Gaps  0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPPDANNIKE 60
        || :||||||||.||||||||:||:.|.|||||  :.:  : |.|:|.|:ILP.|:.:|.|
Db    1 MTVYHFVGIKGTGMSSLAQILHDMKHTVQGSDYEKRFFTQTALEKRNISILPFDKSNVKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
        . |:.:GNAFP:.|| ||.|.: |:||   :.:|| :|||||||||||||||||||||||:
Db   61 GQVIIAGNAFPDTHERIVAAKELNIPVHRYHHFLGDLMNQYTSVAVTGAHGKTSTTGLLA 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||:. : .|:||||||||.||  |::||||||||||||||||||||||||||||||||||:
Db  121 HVMQGAHPTSYLIGDGTGHGVENSKYFVFEACEYRRHFLSYNPDYAIMTNIDFDHPDYFT 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        ||||||.|||||| :||||||| ||||| |:|.| ||:: ||| .:|| ||.||| ..||
Db  181 DINDVFSAFQEMALQVKKGIIACGDDEELQKIQAKVPVIFYGFGEDNDFQARNIQKRTDG 240

Qy  241 TAFDYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        | ||.:.  .| :|:|.:|.:|:.||||||||: .| |. ||.:  |  ||.|||||||
Db  241 TIFDVFVRNTYYDTFKITGYGNHSVLNALAVIALCHYENVDVEAVKHQLTTFEGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
        |   .:.|:|||||||| ||.||| ||:.:|| :||:|||||||||||..|:|.||||||
Db  301 EKPMGEQVIIDDYAHHPTEINATIRAARQKHPEREIVAVFQPHTFSRTEKFLDEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
        |||: .||:|||| ||| |:: |:||:.|| |  |: :.   :|.: :|.|::|FMGAGDI
Db  361 KADQYVLCDIFGSARENKGEJTIEDLQKRIDGAELITDTTTDVLKKKQNGVLIFMGAGDI 420

Qy  421 QKLQNAYLDKL 431
        ||: .||:.:.:
Db  421 QKFEAAYVKEV 431
```

RESULT 9
US-10-282-122A-57306
; Sequence 57306, Application US/10282122A
; Publication No. US20040029129A1
GENERAL INFORMATION:
;   APPLICANT: Wang, Liangsu
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Malone, Cheryl
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Kari
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Forsyth, R.
;   APPLICANT: Xu, H.
;   TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;   FILE REFERENCE: ELITRA.034A
;   CURRENT APPLICATION NUMBER: US/10/282,122A
;   CURRENT FILING DATE: 2003-02-20
;   PRIOR APPLICATION NUMBER: 60/191,078

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10870
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10870

Query Match   63.0%; Score 1433.5; DB 9; Length 456;
Best Local Similarity 60.7%; Pred. No. 1.2e-121;
Matches 258; Conservative 79; Mismatches 87; Indels 1; Gaps 1;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
Db 20 YHFVGIKGSGMSSLALVLHQKGYNVQGSDVEEYFFTQRDLEKSGVPILPFNADNIDKDMI 79

Qy 64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
Db 80 VIAGNAFPDTHEEIARAIELGAEVIRYHDFIARFIEPYTSIAVTGSHGKTSTTGLLAHVL 139

Qy 124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
Db 140 SGINPTSYLIGDGTGHGEPDADPFAFEACEYRRHFLAYSPDYAIMTNIDFDHPDYYKSIE 199

Qy 184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
Db 200 DVFSAFQTMAHQVKKGIFAYGDDKYLRQLESEVPYYYYGVSEEDIQARNIQRTTEGSSF 259

Qy 244 DYYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
Db 260 DVYHKADFVGHFVLPAFGHHNIMNALGVIAVAYFEKLDMQKVABEMLSFKGVKRRFSEKK 319

Qy 304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
Db 320 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFTRTIALMDEFAEALDLAD 379

Qy 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQPDNAVILFMGAGDIQK 422
Db 380 EVFLCNIFGSARETQGEVRIEDLGEKIQKGGQVITEDNVSPLLDFENAEVVFMGAGDVQK 439

Qy 423 LQNAY 427
Db 440 FEQAY 444

RESULT 11
US-10-282-122A-57651
; Sequence 57651, Application US/10282122A
; Publication No. US20040029129A1

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57306
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-57306

Query Match   63.2%; Score 1438.5; DB 12; Length 445;
Best Local Similarity 60.9%; Pred. No. 4.1e-122;
Matches 259; Conservative 79; Mismatches 86; Indels 1; Gaps 1;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
Db   9 YHFVGIKGSGMSSLALVLHQKGYNVQGSDVEEYFFTQRDLEKSGVPILPFNADNIDKDMI 68

Qy 64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
Db 69 VIAGNAFPDTHEEIARAIELGAEVIRYHDFIARFIEPYTSIAVTGSHGKTSTTGLLAHVL 128

Qy 124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
Db 129 SGINPTSYLIGDGTGHGEPDADPFAFEACEYRRHFLAYSPDYAIMTNIDFDHPDYYKSIE 188

Qy 184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
Db 189 DVFSAFQTMAHQVKKGIFAYGDDKYLRQLESEVPYYYYGVSEEDIQARNIQRTTEGSSF 248

Qy 244 DYYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
Db 249 DVYHKDDFVGHFVLPAFGHHNIMNALGVIAVAYFEKLDMQKVAEMLSFKGVKRRFSEKK 308

Qy 304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
Db 309 VSDMIIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFTRTIALMDEFAERALDLAD 368

Qy 364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQPDNAVILFMGAGDIQK 422
Db 369 EVFLCNIFGSARETQGEVRIEDLGEKIQKGGQVITEDNVSPLLDFENAVVFMGAGDVQK 428

Qy 423 LQNAY 427
Db 429 FEQAY 433

RESULT 10
US-09-815-242-10870
; Sequence 10870, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.

Qy 364 RVFLCBIFGSIRENTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDIQK 422
            :|:||:|||:|| :.|||::||||||| |:||:||| .||::||:|::|||||||| :|
Db 368 KVYLCDIFGSAREEQGNVKIEDLGAKIKKGGEVIKENNVSPLLDYHDAVVIFMGAGDVQK 427

Qy 423 LQNAYLDKL 431
            : ||| :||
Db 428 FEQAY-EKL 435

RESULT 12
US-10-282-122A-60988
; Sequence 60988, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60988
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60988

Query Match   59.0%; Score 1342; DB 12; Length 447;
Best Local Similarity   59.7%; Pred. No. 2.5e-113;
Matches 255; Conservative   64; Mismatches 108; Indels   0; Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||:||||||||||| ||.|||| ||| ||||||: |::||:|:|||:||||:||||:| |||:
Db   1 MTIYHFVGIKGSGMSALAQILHDKGFQVQGSDVDKYFFTQKALEBKQIPIMTFSADNIQE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          . |:||| | :::::||| .:|||::::| |||:|:|||||:|||||||||||||||
Db   61 GLTIIAGNAFPDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKTSTTGLLS 120

Qy 121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPYAIMTNIDFDHPDYFK 180

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57651
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57651

Query Match   61.4%; Score 1396; DB 12; Length 444;
Best Local Similarity   60.6%; Pred. No. 3e-118;
Matches 260; Conservative   74; Mismatches   93; Indels   2; Gaps   2;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
          |||||||||||||| |.|||||||::.|: ||||||||:|:||||||||||:|:|:
Db   8 YHFVGIKGSGMSSLALVLHEQGLNVQGSDIEKYFFTQRDLEKANITILPFNADNVKPGMT 67

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
          :| |||| :::::| :.:.|::| |:.| ||||:|||||:|||||:|||||||||||||
Db   68 IIAGNAFPDSHEEIQRAKELGLEVIRYHDFIGHFIQNYTSIAVTGSHGKTSTTGLLSHVL 127

Qy 124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPYAIMTNIDFDHPDYPKDIN 183
          . |: ::|||||:|| |||| :||:||||||||||||| :.||||||||||||||:|:|
Db 128 SGVRPTSYLIGDGTGHGDPQAEFFSFEACEYRRHFLAYSPDYAIMTNIDFDHPDYYTSID 187

Qy 184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIRADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
          ||| |:|| |:|||:|:|| | | :|::| |||:||||:|:.|.||:.:|| |:|.|
Db 188 DVFTAFQTMAGQVKKAIFAYGDDAYLRKLKANVPIYYYGVTENDDIQARNIERTTSGSAF 247

Qy 244 DVYDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
          ||||||| |:| || ||.|:|||:|| | |||:|:|||||:||:.|||||||:
Db 248 DVHGDEFVGHFTVPAFGKHNILNALGVIAVAYFEKLDLKEVAEEMLTFPGVKRRFSEKI 307

Qy 304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
          ||:|::|||||:|| :|:::| |:||:|||..::.|||||||:|:||::|:|||||:||:|
Db 308 VADMTVVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFTRTTALMDEFAEALDLAD 367

```
Query Match           50.1%;  Score 1139;  DB 12;  Length 442;
Best Local Similarity 50.7%;  Pred. No. 7.4e-95;
Matches 218;  Conservative 82;  Mismatches 128;  Indels 2;  Gaps 2;

Qy   4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV  63
        ||:||||||||:| :||||:|  :|||:||||:||:|: ||:  |: :||||  : ||::
Db   5  YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYYFTQRGLEQAGITILPFSEDNITPDME  64

Qy  64  VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTGLLSHVM  123
        :|:|||| ::  :| || | : | :| ||||| |:|| || |:| ||||||||||||:
Db  65  LIVGNAFRENNKEVAYALRHQIPFKRYHDFIGDFMKSFISPAVAGAHGKTSTGLLSHVL  124

Qy 124  NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN  183
        : :  || ||||||| |  |: || |:: :|| ||: |::| | |:|||||||||: |
Db 125  KNITDTSYLIGDGTGRGSANAQYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTGLA  184

Qy 184  DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF  243
        || :||::  : |||| :: |:|:|||||||| ||||||:: .:||:|||  | | :|
Db 185  DVRNAFNDYAKQVKKALFVGEDDBLKKIEAPAPIYYYGFEEGNDFIAYDITRTTNGSDF  244

Qy 244  DVVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT  303
        .|:| | :  : |:|.|| :||||:|||||.|  | :||| |||||::|:|||||||:| |
Db 245  KVKHQGEVIGQFHVPAYGKHNILNATAVIANLFVAGIDMALVADHLKTFSGVKRRFTEKI  304

Qy 304  IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD  363
        |.: .| ||:||||.||.||:  ||:||| ||:|| |||||| ||:| |:|| | | ||
Db 305  INDTIIDDFAHHPTEIVATIDAARQKYPSKEIVAIFQPHTFTRTIALLEDFACALNEAD  364

Qy 364  RVFLCEIFGSIRB-NTGALTIQDLIDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQ  421
        .|:| :|:||:| : :: | ||:|:  ||.    | :|::: |  ||||:|:||||||||
Db 365  SVYLAQIYGSAREVDKGEVKVEDLAAKIIKPSQVVTVENVSPLLDHDNAVVVFMGAGDIQ  424

Qy 422  KLQNAYLDKL  431
        ::::: :: :
Db 425  LYEHSFEELL  434
```

```
Db 121  HVVGAIRPTSYLIGDGTGSGTKDAEYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK  180
        ||: : |:  |||||||| ||||:|||:||||:||||||:|||||:|||||:||||||||
Qy 181  DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG  240
        .|:: || :|:|:||:|||| :|:|:||:|: .:|||:|:|||||.: :|  :.:|: |
Db 181  SVDDVFNAFETLGKQVKKAVFALGDDABLRKLTLDIPIIYFGFGEENEFQAKNVIKETTG  240
Qy 241  TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN  300
        | :|||| : |  | :||.||||:|||||:||||.| || |:||:|::|:|:|||||||
Db 241  TKFDVYHREFLSSFEIPAYGDHNVLNALSVIALCDYEGLPVEDVKNELKTFBGVKRRFS  300
Qy 301  ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS  360
        :|| |.|::|||||||| ||.||:.||:|||| |:||||||||||:||:::||::|||:|
Db 301  ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPDKKVVAVFQPHTFTRTRFLQGFADSLN  360
Qy 361  KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI  420
        :||.||| :|:||:|:.||. |  ||    : .  :: ::.| | ||||:|:|||||||:
Db 361  LADEVYLCDIFGSAREKTGNIFIADLAHKTKGNHIIKEEHTEBLLKYPEAVILFMGAGDV  420

Qy 421  QKLQNAY  427
        ||:| ||
Db 421  QKFQAAY  427
```

RESULT 13

US-10-282-122A-74395

; Sequence 74395, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74395
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74395

RESULT 14

US-10-282-122A-72179

; Sequence 72179, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72179
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72179

```
Query Match             49.8%;  Score 1132.5;  DB 12;  Length 443;
Best Local Similarity   50.7%;  Pred. No. 2.9e-94;
Matches 218;  Conservative 84;  Mismatches 125;  Indels 3;  Gaps 3;

Qy    5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMVV 64
          ||:||||||||| ::| ::|  :|:|||:|:|  |::|    :::|||||||:::|::|
Db    6 HFIGIKGSGMSALALLLHQMGYKVQGSDVDKYFTQHGLEKAGIPILPFAASNITNDMEI 65

Qy   65 IQGNAFASSHE-EIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
         | |||| :     | ||  ::|:: | :||||  : ::||| ||||||||||||| ||:
Db   66 IAGNAFRKDNNIEVAYALENGYHFKRYHEFLGEFMNQFTSLGVAGAHGKTSTTGLLAHVL 125

Qy  124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
          ||:|||||||||| |  ||  :|| | | |:||||  : ||:|||||||||||||:  :
Db  126 KNITDTSYLIGDGTGRGRLANSQYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTGVD 185

Qy  184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
          ||  || : ||:|||| ::||:| ||:| ||||||||||| :||: || :: : ||:|
Db  186 DVFAAFNDYAKQVKKGLFVVGEDPYLRKLTSSAPIYYYGFKDNDDFVAYDIMRSTNGSDF 245

Qy  244 DVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
          :| | ::  : |  |  | |||||:|||:  :|  : |::  :  :||:|||||||:
Db  246 KVRHGQNELGSFHVPAFGRHNVLNATAVIANLYIAGVEMDLVRQHLKTFSGVKRRFSEKL 305

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
          |:: :|:||:||||  |:| ||:||:|||:|:||:|||||||:||: |: ||| :|: |
Db  306 INDVTIIDDFAHHPTEIIATLDAARQKYPSKEIVAIFQPHTFTRTIALLDEFAQALNEAD 365

Qy  364 RVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQPDNAVILFMGAGDIQ 421
          :|: ||:|| :|  : :| :: |: |::   :|:|:: ::::  :|:|||::||||||||
Db  366 SVYLAQIYGSAREVDHGDVKVEDLLADKIVKPAKVITVDNVSPLLDHHNAVYVFMGAGDIQ 425

Qy  422 KLQNAYLDKL 431
          | ::: :  |
Db  426 LYERSFEELL 435
```

RESULT 15
US-09-815-242-13507
; Sequence 13507, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA_011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13507
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13507

```
Query Match             49.2%;  Score 1119;  DB 9;  Length 444;
Best Local Similarity   49.8%;  Pred. No. 4.9e-93;
Matches 213;  Conservative 84;  Mismatches 127;  Indels 4;  Gaps 4;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
          |||:||||||||| ||:|:|:  |:|||||:::|:||  :  |  :|||||||:::|:
Db    5 YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYFFTQRGLEQAGITILPFDEKNLDGDME 64

Qy   64 VIQGNAF-ASSHEBIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV 122
           | |||| :::  :: |  ::|::|:: | :|||: ::: :|  ||||||||||||:|||
Db   65 IIAGNAFRPDNNVEIAYADQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTTGMLSHV 124

Qy  123 MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI 182
          ::  :||||||||||| |   : |||  :  :|:|||  : ||:|||||||||||||:|
Db  125 LSHITDTSFLIGDGTGRGSANAKYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTSL 184

Qy  183 NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFK-DSDDIYAQNIQITDKGT 241
          :|||:|| :  ::: ||:: ||:| :||||:::||||||||   :  ::| :: :: ::
Db  185 EDVFNAFNDYAKQITKGLFVYGEDAELRKITSDAPIYYYGFEAEGNDFVASDLLRSTTGS 244

Qy  242 APDVYVDGEFYDHFLSPQYGDHTVLNALAVAISYLEKLDVTNIKEALETFGGVKRRFNE 301
             | :  ::  : |  |  | ||:||:|||:|   : :::  : |:||:|||||||::
Db  245 TFTVHFRGQNLGQFHIPTFGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTE 304

Qy  302 TTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSK 361
           ::: ::|:||:||||  |:| ||:||:|||:|::||||||||||:||: |: |||: :
Db  305 KIVNDTVIIDDFAHHPTEIIATLDAARQKYPSKEIVAVFQPHTFTRTIALLDDFAHALNQ 364

Qy  362 ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQPDNAVILFMGAGD 419
          ||:|:| :|:|| :|  : :| :: |: :::   ::::: :::::  :|:|||::|||||
Db  365 ADAVYLAQIYGSAREVDHGDVKVEDLANKINKKGHQVITVENVSPLLDHDNAVYVFMGAGD 424

Qy  420 IQKLQNAY 427
          || : :
Db  425 IQTYEYSF 432
```

Search completed: June 3, 2004, 14:46:36
Job time : 50 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    June  3, 2004, 14:38:01 ; Search time 20 Seconds
                              (without alignments)
                              2101.785 Million cell updates/sec

Title:          US-10-712-713-2
Perfect score:  2275
Sequence:       1 MTHYHFVGIKGSGMSSLAQI.........GDIQKLQNAYLDKLGMKNAP 437
Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters:      283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     PIR 78:*
               1: pir1:*
               2: pir2:*
               3: pir3:*
               4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2274	100.0	437	2	H89958	UDP-N-Aacerylmuram
2	1438	63.2	432	2	C69662	UDP-N-acetylmurama
3	1368	60.1	433	2	H84055	UDP-N-acetylmurama
4	1355	59.6	447	2	AE1638	UDP-N-acetyl muram
5	1342	59.0	447	2	AE1275	UDP-N-acetyl muram
6	1180	51.9	457	2	A86879	hypothetical prote
7	1119	49.2	444	2	D98043	UDP-N-acetylmurama
8	1117	49.1	444	2	B95177	UDP-N-acetylmurama
9	641.5	28.2	458	2	F97296	UDP-n-acetylmurama
10	566	24.9	495	2	C71679	UDP-n-acetylmurama
11	565.5	24.9	485	2	C97741	hypothetical prote
12	548	24.1	506	2	A81777	UDP-N-acetylmurama
13	538	23.6	468	2	H70201	UDP-N-acetylmurama
14	535	23.5	469	2	B81201	UDP-N-acetylmurama
15	534	23.5	473	2	A87565	UDP-N-acetylmurama
16	521	22.9	454	2	B70418	UDP-N-acetylmurama
17	513	22.5	471	2	AF2833	UDP-N-acetylmurama
18	513	22.5	471	2	B97611	UDP-N-acetylmurama
19	500.5	22.0	494	2	AI2438	UDP-N-acetylmuramo
20	496	21.8	833	2	B81737	UDP-N-acetylmurama
21	492.5	21.6	471	2	AF3324	UDP-N-acetylmurama
22	491	21.6	491	1	CEECAM	UDP-N-acetyl-muram
23	491	21.6	491	2	G90640	UDP-N-acetyl-muram
24	491	21.6	491	2	G85491	UDP-N-acetyl-muram
25	490.5	21.5	491	2	AF0518	UDP-N-acetylmurama
26	490	21.5	480	2	D83094	UDP-N-acetylmurama
27	487	21.4	481	2	C71338	probable UDP-N-ace
28	470	20.7	809	2	G86603	muramate-Ala ligas
29	470	20.7	809	2	D72022	UDP-N-acetylmurama
30	465	20.4		2	AI0068	UDP-N-acetylmurama
31	463	20.4		2	A72402	UDP-N-acetylmurama
32	463	20.4	803	2	A71475	probable muramate-
33	461.5	20.3	462	2	T35852	probable UDP-N-ace
34	460.5	20.2	475	2	B64185	UDP-N-acetylmurama
35	460.5	20.2	506	2	D82081	UDP-N-acetylmurama
36	455	20.0	477	2	E82763	UDP-N-acetylmurama
37	453.5	19.9	432	2	C81308	UDP-N-acetylmurama
38	446	19.6	484	2	C84955	UDP-N-acetylmurama
39	442.5	19.5	505	2	S76722	UDP-N-acetylmuramo
40	420	18.5	494	2	D70579	probable murC prot
41	417.5	18.4	495	2	E87023	UDP-N-acetyl-muram
42	410	18.0	449	2	B71917	udp-n-acetylmurama
43	403	17.7	449	2	G64597	UDP-N-acetylmurama
44	392.5	17.3	453	2	B64002	UDP-N-acetylmurama
45	390	17.1	458	2	F81904	UDP-N-acetylmurama

```
                         ALIGNMENTS

RESULT 1
H89958
UDP-N-Aacerylmuramate-alanine ligase [imported] - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: H89958
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: H89958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-437 <KUR>
A;Cross-references: GB:BA000018; PID:BA000018; PID:g13701535; PIDN:BAB42829.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match             100.0%;  Score 2274;  DB 2;  Length 437;
Best Local Similarity   99.8%;  Pred. No. 1.4e-148;
Matches  436;  Conservative  1;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKILPPDANNIKE 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKILPPDANNIKE 60

Qy     61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 DMVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGKDSDDIYAQNIQITDKG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGKDSDDIYAQNIQITDKG 240

Qy    241 TAPDVYVDGEFYDHFLSPQYGDHTVLNALAVAISYLEKLDVTNIKEALETFGGVKRRFN 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAPDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFABSLS 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFABSLS 360

Qy    361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db 361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420

Qy 421 QKLQNAYLDKLGMKQNAF 437
            ||||||||||||||||
Db 421 QKLQNAYLDKLGMKQNAF 437

RESULT 2
C69662
UDP-N-acetylmuramate-alanine ligase (EC 6.3.2.8) murC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 03-Jun-2002
C;Accession: C69662; S71002
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero,
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho-
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee;
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A.;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.;
A.;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <KUN>
A;Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB14957.1; PID:g2635463
R;Varon, D.; Brody, M.S.; Price, C.W.
Mol. Microbiol. 20, 339-350, 1996
A;Title: Bacillus subtilis operon under the dual control of the general stress transcrip
A;Reference number: S71000; MUID:96310371; PMID:8733232
A;Accesion: S71002
A;Molecule type: DNA
A;Residues: 85-432 <VAR>
A;Cross-references: EMBL:L31845; NID:g556013; PIDN:AAB40043.1; PID:g556014
A;Experimental source: strain 168, substrain Marburg
C;Genetics:
A;Gene: murC
C;Function:
C;Pathway: peptidoglycan biosynthesis
C;Superfamily: UDP-N-acetylmuramate-alanine ligase
C;Keywords: ATP binding; cell division; cell wall; ligase; peptidoglycan biosynthesis

Query Match   63.2%; Score 1438; DB 2; Length 432;
Best Local Similarity   64.6%; Pred. No. 3.4e-91;
Matches 276; Conservative 51; Mismatches 100; Indels 0; Gaps 0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          | |||:||:|||:||| |::: | |::||||: :|| | | | ||| ||
Db   1 MTVVHFVGIKGTGMSPLAQILHDNGYTVQGSDIEKFIFTQTALEKRNITILPFSAENIKP 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          ||: |||||| :| | |:| |||::||| |:|:: ||||:|||||||||||||
Db   61 GMTVIAGNAFPDTHPEIEKAMSEGIPVIRYHKFLGDYMKKFTSVAVTGAHGKTSTTGLLA 120

Qy 121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFHDPDYFK 180
          ||:| | || |||||||| | | | | |:||| |||:: :|||||||||||:
Db 121 HVIQNAKPTSFLIGDGTGQGQNENSRYFVFEACEYRRHFLSYQPDYAIMTNIDFHDPDYFS 180

Qy 181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ||||||||| :: || |:| ||||||: | || |:|:: ||: | |||::| |
Db 181 SIDDVFDAFQEMALQVNKGIIACGDDEHLPKIHANVPVVYVGTGENDFQARNIVKSTEG 240

Qy 241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          ||||||:|||||||||||||:|||||||||||||| ||: |||| |||||:|||:||:
Db 241 (continues)

RESULT 3
H84055
UDP-N-acetylmuramate-alanine ligase murC [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84055
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84055
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-433 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06967.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   60.1%; Score 1368; DB 2; Length 433;
Best Local Similarity   59.3%; Pred. No. 2.2e-86;
Matches 253; Conservative 69; Mismatches 105; Indels 0; Gaps 0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          | :||:||||||||||| |: |||::|:|:|||| |::|| | |||:| || |||
Db   1 MTFHFIGIKGSGMSALAQVLHDMNYQVQGSDVEKRFFTQKPLEEKGIPIPFQKENIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          |:: :: | :| |: |: | |:| || :| | | |||:||| ||||||||
Db   61 GQQVVVSAAYSDEHEEVVAAKEQGLEIQSYPTFLGBFIQRFTSIAVTGSHGKTSTTGLLS 120

Qy 121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFHDPDYFK 180
          ||: ||||||||:| | |||: | : | |||||:| :|||||||||||| |
Db 121 HVLGSHCPTSFLIGDGTGKGAADSHYFVFEACEYRRHFLNYRPDYCIMTNIDFDHPDYFD 180

Qy 181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ||||||:|||||| || ||:| | :: |: |: |: : |:: | |||:|| |
Db 181 DINDVFSAFQEMAMQVKAIVACGDDEQLQGIHANVPVVVGFNEDNDFQARNIEVTSEG 240

Qy 241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          |::|:|:|||||||:|:|| ::||| ||||| |||::| |||| |||| :|||:||:|:|
Db 241 THFQVFVRNNHYGDFTIPRFGKHNVLNSLAVIALCHYENIPVEVLARELKTFEGVKRRFT 300

Qy 301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHFTSRTQAFLNEFAESLS 360
          |::|:||| |||||||| :|||||| |||:|| |::| ||||:||||| ||||| |:
Db 301 EKTLGGQILIDDYAHHPTEISAATTEEAAKQKYPDRPVAIFQPHTFSRTKAFLHEFAEALK 360

Qy 361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          :||:|:|||||| :||| || : ||| |: |:||::|:: | |:||:::|||||||
Db 361 QADHVYLCDIFGSARENKGSLSIEDLQALIPGAHLLSENDLAPLQDHRSGVLLFMGAGDI 420

Qy 421 QKLQNAY 427
           || | |
Db 421 QKFQQAY 427

RESULT 4

```
AE1638
UDP-N-acetyl muramate-alanine ligases homolog murC [imported] - Listeria innocua (strain
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 14-Dec-2001
C;Accession: AE1638
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96877.1; PID:g16414133; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match           59.6%;    Score 1355;   DB 2;    Length 447;
Best Local Similarity 60.2%;    Pred. No. 1.8e-85;
Matches  257; Conservative  63; Mismatches 107;    Indels   0;   Gaps   0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||:|||||||||||| ||||:||: |::||||||:  :|:|:|||||||||| :
Db    1 MTIYHFVGIKGSGMSALAQILHDKGFQVQGSDVDKYFFTQKALEEKQIPIMTFSADNIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          |: :|||||||  ||||  | : |||:|:|:|||||:||||||::|:||||||||||||
Db   61 GLTIIAGNAFPDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||:    :|:||||||| || |||||||:|||||:|||| |||| ||||||:  |||||
Db  121 HVVGAIRPTSYLIGDGTGSGTKDAYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ::|||:  |:  :|||:|:::|  |:|||:| | |:||| :||:: :   |:
Db  181 SVDDVFNAFETLGKQVKKAVFALGDDVELRKLSLDVPIIYFGFGEENEFQAKNVKKETTG 240

Qy  241 TAFDYVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          | ||:| ||||:|:||| :|:|||:|||||| | |:|:|| || || |:|||||||
Db  241 TKFDVYHREFLASFEIPAYGDHNVLNALSVIALCDYEGLPVEDVKELKTFEGVKRRFS 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          ||  ||||:||||||||:||:|| ||:|||| ::||||||||||:|| |||:  |:||
Db  301 ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPDKKVVAVFQPHTFTRTRTFLQGFADSLN 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKEGASLINEDSINVLEQFDNAVLFMGAGDI 420
          ||| ||| :||||:|:|||  || |:| || |:||:|:::  | :| |||||||||:
Db  361 LADEVYLCDIFGSAREKTGNLTIADLAHKTKGNHIIKEEHTEELLQYPEAVLFMGAGDV 420

Qy  421 QKLQNAY 427
          || ||
Db  421 QKFQAAY 427
```

```
RESULT 5
AE1275
UDP-N-acetyl muramate-alanine ligases homolog murC [imported] - Listeria monocytogenes
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 14-Dec-2001
C;Accession: AE1275
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
```

```
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1275
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99683.1; PID:g16411034; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match           59.0%;    Score 1342;   DB 2;    Length 447;
Best Local Similarity 59.7%;    Pred. No. 1.4e-84;
Matches  255; Conservative  64; Mismatches 108;    Indels   0;   Gaps   0;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||:|||||||||||| ||||:||: |::||||||:  :|:|:|||||||||| :
Db    1 MTIYHFVGIKGSGMSALAQILHDKGFQVQGSDVDKYFFTQKALEEKQIPIMTFSADNIQE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          |: :|||||||  ||||  | : |||:|:|:|||||:||||||::|:||||||||||||
Db   61 GLTIIAGNAFPDTHEEIERALELGLSVIRYHKFLGQLIDGYTSIAITGSHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||:    :|:||||||| || |||||||:|||||:|||| |||| ||||||:  |||||
Db  121 HVVGAIRPTSYLIGDGTGSGTKDAYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ::|||:  |:  :|||:|:::|  |:|||:| | |:||| :||:: :   |:
Db  181 SVDDVFNAFETLGKQVKKAVFALGDDAELRKLTLDIPIIYFGFGEENEFQAKNVIKETTG 240

Qy  241 TAFDYVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          | ||:| ||||:|:||| :|:|||:|||||| | |:|:|| || || |:|||||||
Db  241 TKFDVYHREFLSSFEIPAYGDHNVLNALSVIALCDYEGLPVEDVKNELKTFEGVKRRFS 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          ||  ||||:||||||||:||:|| ||:|||| ::||||||||||:|| |||:  |:||
Db  301 ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPDKKVVAVFQPHTFTRTRTFLQGFADSLN 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKEGASLINEDSINVLEQFDNAVLFMGAGDI 420
          ||| ||| :||||:|:|||  || |:| || |:||:|:::  | :| |||||||||:
Db  361 LADEVYLCDIFGSAREKTGNLTIADLAHKTKGNHIIKEEHTEELLKYPEAVLFMGAGDV 420

Qy  421 QKLQNAY 427
          || ||
Db  421 QKFQAAY 427
```

```
RESULT 6
H86879
hypothetical protein murC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: H86879
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86879
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <STO>
A;Cross-references: GB:AE005176; PID:g12725089; PIDN:AAK06138.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match           51.9%;    Score 1180;   DB 2;    Length 457;
Best Local Similarity 52.7%;    Pred. No. 1.9e-73;
Matches  227; Conservative  79; Mismatches 121;    Indels   4;   Gaps   4;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
```

```
Db   185  EDVFNAFNDYAKQITKGLFVYGEDALRKITSDAPIYYGFRAEGNDFVASDLLRSTTGS  244
          ||:|: . :|| |||||||:|  :|||
Qy   242  AFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNE  301
          |:|:. |  :  |: .||||:|  . .:|:|
Db   245  TFTVHFRGQNLGQFHIPTFGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTE  304
          |||||||:|
Qy   302  TTIANQVIVDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSK  361
          |||||||:||||||||||||||||||||||||| ||||||:|||||||||||||:|||
Db   305  KIVNDTVIIDDFAHHPTEIIATLDAARQKYPSKEIVAVFQPHTFRTIALLDDFAHALNQ  364
          ||:|||:||||||:||: .::|| : :|:.:
Qy   362  ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGD  419
          |:|||| :|||||||   ||||||||||||  ||||||||:||||:||:|||||||||||
Db   365  ADAVYLAQIYGSAREVDHGDVKVEDLANKINKKHQVITVENVSPLLDHDNAVVVFMGAGD  424

Qy   420  IQKLQNAY  427
          | :|.: :
Db   425  IQTYEYSF  432
```

RESULT 8
B95177
UDP-N-acetylmuramate-alanine ligase [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: B95177
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
nson, T.; Hickey, E.K.; Holt, I.E.
J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: B95177
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75611.1; PID:g14973011; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1521
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

```
Query Match              49.1%;  Score 1117;  DB 2;  Length 444;
Best Local Similarity    49.8%;  Pred. No. 3.9e-69;
  Matches   213;  Conservative   84;  Mismatches  127;  Indels    4;  Gaps    4;

Qy     4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPPDANNIKEDMV  63
          ||:|||||||||||||:|||||||||:|||||||:|:||||||||||||||||||||||
Db     5  YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYYFTQRGLEQAGITILPFDEKNLDGDME  64

Qy    64  VIQGNAF-ASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV  122
          :|:||:    ||::|:||:| :.|||||:| |:.|.||:|.||:|||||||||||:|||
Db    65  IIAGNAFRPDNNVEIAYADQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTTGMLSHV  124

Qy   123  MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI  182
          ::  ||||||||||||||||||||:||:: :|||||||||||||||:||||||||||||:
Db   125  LSHITDTSFLIGDGTGRGSANAKYFVFESDEYERHFMPYIHPEYSIITNIDFDHPDYFTSL  184

Qy   183  NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFK-DSDDIYAQNIQITDKGT  241
          ||:|: . :|| |:.|| || ||:::|||||||||:|||   ||: :|  |||:|||||
Db   185  EDVFNAFNDYAKQITKGLFVYGEDALRKITSDAPIYYGFRAEGNDFVASDLLRSTTGS  244

Qy   242  AFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFTE  301
          |:|:. |  :  |: .||||:|  . .:|:|
Db   245  TFTVHFRGQNLGQFHIPTFGRHNIMNATAVIGLLYTAGFDLNLVREHLKTFAGVKRRFTE  304

Qy   302  TTIANQVIVDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSK  361
          |||||||:||||||||||||||||||||||||| ||||||:|||||||||||||:|||
Db   305  KIVNDTVIIDDFAHHPTEIIATLDAARQKYPSKEIVAVFQPHTFRTIALLDDFAHALNQ  364

Qy   362  ADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGD  419
          |:|||| :|||||||   ||||||||||||  ||||||||:||||:||:|||||||||||
Db   365  ADAVYLAQIYGSAREVDHGDVKVEDLANKINKKHQVITVENVSPLLDHDNAVYVFMGAGD  424
```

```
Db    19  YHFTGIKGSGMSALALMLHQMGKKVQGSDTDYFFTQRGLEQADVPLLPFDEKNIKEFE  78
          |||||||||||||:|||||||:||:.  : : |||:||||||  :|| |||:|||
Qy    64  VIQGNAFASSHE-EIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV  122
          :|:||:|  ..| : |.|   | :||:.|:| :  : |. ||:|||||||||||:|||
Db    79  LIAGNAFRDDNNVEIAFAHKNGFPFKRYHEFLGHFMEDFTSIGVAGAHGKTSTTGMLAHV  138

Qy   123  MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI  182
          ::  ||||:|||||| :|||||||:||:: :|||||:||  |:  ||:||||||||:||
Db   139  MSNIVDTSYLIGDGTGRGIEGSSEYFVFESDEYERHFMPYHPEYTIMTNIDFDHPDYFGI  198

Qy   183  NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKGTA  242
          ||:|::. :|| ||:|| ||:||:|:|||:|||||:|| |||:: || || |||  :
Db   199  EDVTSAFQDYANNIKKGIFAYGEDVNLRKLTAKAPIYYGFEANDDYRAENLVRSTRGSS  258

Qy   243  FDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNET  302
          ||.: |  |::|:   |||:|   .: |:| .|: .|| :.:. |.   :|:|||||:|
Db   259  FDAYFRGEKIGHFVVPAYGKHNVLNALSVVAVCHNLGLDMTEVADHLLTFRGVKRRFTEK  318

Qy   303  TIANQVIVDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKA  362
          ||||||:||||||||||||||||||||||| ||||||:|||||||||||||:|:| |
Db   319  KVGETVIIDDFAHHPTEIEATLDAARQKYPDREIVAVFQPHTFRTIAFADEFAEVLDHA  378

Qy   363  DRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDI  420
          | ||| :|||||||   ||||||||||||  ||||||||:||||:|||||:| ||:.|||
Db   379  DTVYLAQIYGSAREVDHEITAQDLADKVRKPAKVIDLDNVSPLLDHDRGVYVVFMGAGNI  438

Qy   421  QKLQNAYLDKL  431
          |:|||| :|||
Db   439  QKYELAF-EKL  448
```

RESULT 7
D98043
UDP-N-acetylmuramate-alanine ligase (EC 6.3.2.8) [imported] - Streptococcus pneumoniae
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
C;Accession: D98043
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; Sun, P.M.; Winkler, M.E.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D98043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00177.1; PID:g15459022; GSPDB:GN00174
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase
C;Keywords: ligase

```
Query Match              49.2%;  Score 1119;  DB 2;  Length 444;
Best Local Similarity    49.8%;  Pred. No. 2.8e-69;
  Matches   213;  Conservative   84;  Mismatches  127;  Indels    4;  Gaps    4;

Qy     4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPPDANNIKEDMV  63
          ||:|||||||||||||:|||||||:|||||||:|:||||||||||||||||||||||||
Db     5  YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYYFTQRGLEQAGITILPFDEKNLDGDME  64

Qy    64  VIQGNAF-ASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV  122
          :|:||:    ||::|:||:| :.|||||:| |:.|.||:|.||:|||||||||||:|||
Db    65  IIAGNAFRPDNNVEIAYADQNGISYKRYHEFLGSFMRDFVSMGVAGAHGKTSTTGMLSHV  124

Qy   123  MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI  182
          ::  ||||||||||||||||||||:||:: :|||||||||||||||:||||||||||||:
Db   125  LSHITDTSFLIGDGTGRGSANAKYFVFESDEYERHFMPYIHPEYSIITNIDFDHPDYFTSL  184

Qy   183  NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFK-DSDDIYAQNIQITDKGT  241
          ||:|: . :|| |:.|| || ||:::|||||||||:|||   ||: :|  |||:|||||
Db   185  EDVFNAFNDYAKQITKGLFVYGEDALRKITSDAPIYYGFRAEGNDFVASDLLRSTTGS  244
```

Qy   420 IQKLQONAY 427
           | | : ::
Db   425 IQTYEYSF 432

RESULT 9
F97296
UDP-N-acetylmuramate-alanine ligase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 30-Sep-2001
C;Accession: F97296
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: F97296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-458 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK81161.1; PID:g15026298; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3225
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   28.2%;   Score 641.5; DB 2; Length 458;
Best Local Similarity 33.5%; Pred. No. 1.8e-36;
Matches 150; Conservative 108; Mismatches 165; Indels 25; Gaps 16;

Qy   5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKEDMV 63
          ||:|| ||||||||||||||||:|||:|| :|||:||||||||||| :||||||||||
Db   14 HFIGIGGVGMSMGLAEILLERGFKVSGGDMNGSPMID-KLKEHGAEIYLGHNEKNINNVDI 72

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYT-SVAVTGAHGKTSTGLLSHV 122
          |||||||||||||||||||||:|||||||||||||::|| :|||||||||||:||:|||
Db   73 VVYTAAIPEDNPELIYARKNNISLMTRAEFLGSLMKGHKYNIAISGTHGKTTTSMVSHI 132

Qy   123 -MNGDKKTSFLIG------DGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
           :|||||||||| ||:||||||||||||:||||||| :||| ||| |||||||
Db   133 ALTEDVDPTTIVGGNLDIINGNVLA-GKSDYFITEACEYKASFLEFYPIYGIVILNIDADH 191

Qy   176 PDYFKDINDVFDAFQEMAHNV-KKG-IIAWGDDEHLRKIEADVPIYYYGFK-DSDDIYAQ 232
          |||||||||||||||||| | || :|:||||||||||:::||:||::| :| ||::
Db   192 LDYYKNIDDIRNTFAKFVNLIPKEGYLIANADDKRVARVASNATCNVVSFGIDNGDIRAK 251

Qy   233 NIQITDKG-TAFDVYVDGEFYDHFLSPQYGDHTVNLALAVIAISYLEKLDVTNIKEALET 291
          :| |:| |||||||| ||||||:| ::||||:|||:|||:|||:|||:||||||||
Db   252 NISFNESGFSSFDVYKSSELLFNIELNVPGKHNILNALSAIASALTLKISHKSIIDGLKS 311

Qy   292 FGGVKRRFNETTIANQV-IVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQA 350
          |||||||||||||||| :||||||||||||||||||||| :||||||||:|:|||
Db   312 FKGTHRRFEIKGVKNGITVIDDYAHHPEIKATLDAA-KNYPHNKIYCVFQPHTYSRTLS 370

Qy   351 FLNEFAESLSKADRVFLCEIFGSIRENTG--ALTIQDLIDK--IEGASLIN-EDSINVL 404
          |::|||||||||||||||||||||||| : :|||| |:|:   ::|:| | ||::|
Db   371 LFDDFSNSFSGVDELVLADIYAAREKDTGVVSSLKLSEAINKNGVKSSNLHSFEDIVNYF 430

Qy   405 EQ--FDNAVILFMGAGDIQKLQNAYLDK 430
          || :|| ||||||:||:|:||:||
Db   431 KSKLNDGDILLTVGAGDVFKIGEMFLSK 458

RESULT 10
C71679
UDP-n-acetylmuramate-alanine ligase (murC) RP247 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C;Accession: C71679
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: C71679
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-495 <AND>
A;Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14709.1; PID:g386080
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: murC; RP247
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   24.9%;   Score 566; DB 2; Length 495;
Best Local Similarity 33.3%; Pred. No. 3e-31;
Matches 158; Conservative 78; Mismatches 172; Indels 66; Gaps 17;

Qy   5 HFVGIKGSGMSSLAQIMHDLGHEVQGSD-IENYVFTEVALRNKGIKILPFDA-NNIKEDM 62
          ||:|| |||||||||||||||||||:|| :||:|||||||||||||||||:| :||:|:|
Db   17 HFIGIGGVGMSGIARTLHNLGYKVQGSDLVENY--NTKRLESYGIKIFLGQAKQNIKNVS 74

Qy   63 VVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTT------ 116
          ||:|||||||||||||||||||:||||||||||||:::::|||||||||||:||
Db   75 YVVISSAINQNNPEIKEALERKIPIIRRAEMLAELMRLKCSVAVSGSHGKTTTSLIACL 134

Qy   117 ---GLLSHVMNG----DKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIM 168
             |||:|: | : |:||||||||||||:||:||||:||||||||||:|||||:
Db   135 FEAAGLCPTVINGGIINNKSTNAYLG------SSNYLIAEADESDATFIHIPSTIAII 186

Qy   169 TNIDFDHPDYFKDINDVFDAFQEMAHNVKKGIIAWG-------DDEHLRKIEADV---PIY 219
          |||| |:||||||||:|||:|| :| | || || : :::||||:|||:|| ||:
Db   187 TNIDPEHLDYYQDFBILIGAFRSFITNLP---FYGFAVCCIDHKIVRKLVDDITERKI 242

Qy   220 YYGFKDSDD--IYAQNIQITDKGTAFDVYVD----------GEFYDHFLSPQYGDHTVLNALAV 271
          ||| | || :|:: ||| |||||||||| | ||||||||||:||||:|||||:|
Db   243 TYGI-DSEDAHIIAFNINTDIASSTFDVKISLPNVLGTTIEKITIPTPGRHNILNSLAA 301

Qy   272 IAISYLEKLDVTNIKEALETFGGVKRRFNETTIANQ-VIVDDYAHHPREISATIETARK- 329
          |||||:| ||:||||||:||:||||||||||||||| :||||||||||||:||:|:: |:
Db   302 IAVGIELDFGIKAIKNGFNNFKGVKRRFTKVAEYNKAVIIDDYAHHPEIKATLATAKNI 361

Qy   330 -KYPHKEVVAVFQPHTPSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLID 388
           ::|| |||| |||||:||||||||||||:||||||:||:|||||:|:|||:||||:|||
Db   362 ANQQNGKVIAIFQPHRYSRIKYLFDDFMLCFADADILYITNIYAAGETPIEGITGQSLVD 421

Qy   389 KIEGASLINEDSINVLEQFDNAV----------------ILFMGAGDIQKLQNAYLDK 430
          ::|:|| ||:|||:|:|:|||| :|:|:||:||:|||:||
Db   422 KMTQNK--HHDKANFLAELDVVSIIDNAASGDMIIMMGAGNISSPANE-LDR 472

RESULT 11
C97741
hypothetical protein murC [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: C97741
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RO
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: C97741
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-485 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL02869.1; PID:g15619392; GSPDB:GN00173
C;Genetics:
A;Gene: murC
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   24.9%;   Score 565.5; DB 2; Length 485;
Best Local Similarity 33.5%; Pred. No. 3.2e-31;
Matches 161; Conservative 76; Mismatches 173; Indels 71; Gaps 18;

Qy   5 HFVGIKGSGMSSLAQIMHDLGHEVQGSD-IENYVFTEVALRNKGIKI-LPFDANNIKEDM 62
          ||:|| |||||||||||||||||||:|| :||:|||||||||||||| |

```
Db   17  HFIGIGGVGMSGIAEILYNLGYKVQGSDLVENY--NTKRLESYGIKIFLGHAEQNITNVS  74
Qy   63  VVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTT------  116
Db   75  YVVISSAINPKNPEIKEALERKIPIIRRADMLAELMLRKCSVAVSGSHGKTTTTSLVACL  134
Qy  117  ---GLLSHVMNG----DKKTSFLIGDGTGMGLPESDYFAFFACEYRRHFLSYKPDYAIM   168
Db  135  FEAAGLCPTVINGGIINNKSTNAYLG------SSNYLIAEADESDATFIHIPSTIAII    186
Qy  169  TNIDFDHPDYFKDINDVFDAFQEMAHNVKKGIIAWG------DDEHLRKIEADV---PIY  219
Db  187  TNIDPEHLDYYRDFETLIGAFRSFIITNLP----FYGFAVCCIDHKIVRKLVDDITERKIV  242
Qy  220  YYGFKDSDD--IYAQNIQITDKGTAFDVYVD-----GEFYDHFLSPQYGDHTVLNALAV   271
Db  243  TYGI-DSEDAHIIAFNINTDIASSTFDVKISLPNVLGTTIIEKITIPTPGRHNILNSLAA  301
Qy  272  IAISYLEKLDVTNIKEALETFGGVKRFNETTIA--NQVIVDDYAHHPREISATIETAR    328
Db  302  IAVGIELDFGIKAIKNGFNNFKGVKRRF--TKVAEYNNASIIDDYAHHPEEIKATLATA-  358
Qy  329  KKYPHKE---VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQD  385
Db  359  KNIANKQNGKVIAIFQPHRYSRMQYLFFDDFMLCFADADILYITDIYAAGENPIEGITGRS  418
Qy  386  LIDKIEGASLINEDSINVLEQFDNAV---------ILFMGAGDIQKLQNAYLDKLGM     433
Db  419  LVDKITKRK--HHDKANFLAEILDDAVGVIIDNAASGDMIIMWGAGNISSFANELEGRLSS  476
Qy  434  K  434
Db  477  R  477
```

RESULT 12
A81777
UDP-N-acetylmuramate-alanine ligase (EC 6.3.2.8) NMA2061 [imported] - Neisseria meningit
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: A81777
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-506 <PAR>
A;Cross-references: GB:AL162758; GB:AL157959; NID:g73380672; PIDN:CAB85279.1; PID:g738068
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: murC; NMA2061
C;Superfamily: UDP-N-acetylmuramate-alanine ligase
C;Keywords: ligase

```
Query Match         24.1%;     Score 548;   DB 2;   Length 506;
Best Local Similarity   32.5%;     Pred. No. 5.4e-30;
Matches 151;   Conservative  79;   Mismatches 193;   Indels 42;   Gaps 14;

Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNIK  59
Db   43  VTNIHFVGIGGVGMSGIAEVLHNLGFKVSGSDQARNAATE-HLGSLGIQVTPGHTAEHVN  101
Qy   60  EDMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLL  119
Db  102  GADVVVTSTAVKKENPEVVAALEQQIPVIPRALMLABLMRFRDGIAIAGTHGKTTTTSLT  161
Qy  120  SHVMN-GDKKTSFLIGD-----GTGMGLPESDYFAFFACEYRRHFLSYKPDYAIMTNIDF  173
Db  162  ASILGAAGLDPTFVIGGKLNAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDE  221
```

```
Qy  174  DHPD-YFKDINDVFDAFQEMAHNVK---KGIIAWGDDEHLRKI--EADVPIYYYGFKDSD  227
Db  222  DHMDYDYGHSVEKLHQAFIDFIHRMPFYGKAFICI-DSEHVRAILPKVSKPYATYGLDDTA  280
Qy  228  DIYAQNIQITDKGTAPDVY-----------DGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLD  281
Db  281  DIYATDIENVGAQMKFTVHVQMKGHEQGSFEVVLNMP--GRHNVLNALLAAIGVALEVGAS  338
Qy  282  VTNIKEALETFPGGVKRF---NETTIAN--QVIVDDYAHHPREISATIETARKKYPHKE  335
Db  339  VEAIQKGLGLLGFEGVGRRFQKYGDIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKR  398
Qy  336  VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKI-----  390
Db  399  LVLAFQPHRYTRTRDLPEDFTKVLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGK  458
Qy  391  ------EGASLINEDSINVLEQFDNAVILFMGAGDIQKLQNAYLD  429
Db  459  LEPIYCENVADLPEMLLNVLQ--DGDIVLNMGAGSINRVPAALLE  501
```

RESULT 13
H70201
UDP-N-acetylmuramate-alanine ligase (murC) homolog - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: H70201
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, S.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: H70201
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-468 <KLE>
A;Cross-references: GB:AE001180; GB:AE000783; NID:g2688755; PIDN:AAC67166.1; PID:g268876
A;Experimental source: strain B31
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

```
Query Match         23.6%;     Score 538;   DB 2;   Length 468;
Best Local Similarity   31.0%;     Pred. No. 2.3e-29;
Matches 148;   Conservative  91;   Mismatches 155;   Indels 84;   Gaps 19;

Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  60
Db    8  LNNIFFVGIKGSGGACSLACFLNSKGYCVEGVDVSDKFYTDEILSNN--KISYYD--NIYE  63
Qy   61  ----------DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAH  110
Db   64  FSLKQLDRSFDLIVYSSAYNKDGLQVLLEAKELNIPILSYPEALGELSRKYYSIGIAGSH  123
Qy  111  GKTSTT----------GLLSHVMNGDKKTSFLIGDGTGM-GLPESDYFAFFACEYRRHFL  159
Db  124  GKTTTTAFLGVLFNKLGLNPNVIVGSSVKDF--KDNSAIAGI--SNIFIVETCEYKKHFL  179
Qy  160  SYKPDYAIMTNIDFDHPDYFKDINDVFDAFQEMAHNVKKG--IIAWGDDEHLRKIE----  213
Db  180  NFSPNLLILTNVDYEHVDFFKNYEALEEAFLQYINNLKKNGILINSDDNNLLKIKRQIN  239
Qy  214  -ADVPIYYYGFKDSDDIYAQNIQITDKGTAFDVYVDGEFYDHFLSPQYGD------HTVL  266
Db  240  RKDISIFSYGSGDLSDFQISNIAVRSEYFCFS------------FLGLLNVELKTVLFHNVL  289
Qy  267  N-ALAVIAIS-YLEK--LDVTNIKEAL----ETFGGVKRFNETTIANQVI-VDDYAHHP  317
Db  290  NFSAALLALNLFLESNGKSIFDFEEAIKRIAKNYSGIKRVEVVKENGVIYMDDYAHHP  349
Qy  318  REISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIREN  377
Db  350  REIKNTLFGIKNFYKNKRIILDFMPHTFTRTKEFFADFVEVLSAADILILHNIYLSNREN  409
```

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>
A;Cross-references: GB:AE005673; NID:g13424109; PIDN:AAK24517.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2546
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   23.5%;   Score 534;   DB 2;   Length 473;
Best Local Similarity   32.0%;   Pred. No. 4.4e-29;
Matches 143; Conservative 82; Mismatches 186; Indels 36; Gaps 12;

Qy   5 HPFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKEDMV 63
            ||:||.|||| |:|:||||||||||||||| |:.: |||| |:.::
Db   16 HFIGIGGIGMSGIAEIMIRIGYTVQGSDAKASANTE-RLEKLGARIFIGHDAAHVEGASA 74

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVSYNDFLGQIIDQYTSVAVTGAHGKTSTT------- 116
            :: ||||||||||||||||||||||||||||||||||:|:|||||||||||:::
Db   75 IVYSTAVKADNPEMVAGRDKRLPLVRRAEMLABLMRLQFSVAVGGTHGKTTTTSMVAALL 134

Qy   117 ---GLLSHVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDF 173
                    |:|:| :|| ||:|:||||
Db   135 DAGGLDPTVVNGGIINAY----GTNAKVGEGDWIVEADESDGSFLKLKSTVAIVTNIDA 190

Qy   174 DHPDYFKDINDVFDAFQEMAHNVK--KGIIA-WGDDEHLRKIRADVP---IYYYGFKDSDD 228
            ||||||||:||||||||||||||| ||||. :|:: :|:||| : .::|:||.|||:
Db   191 EHLDHWGDFDAVKKGFQDFIQNIPFYGFAAVCTDHPEVQALTSRIENRRLVTYGTNPQAE 250

Qy   229 IYAQNIQITDKGTAFDYVD---GEF--YDHFLSPQYGDHTVLNALAVIAISYLEKLDVT 283
            :.|:||||:|||||||:|. |:| ||:|| ||||||:: |||||:|:||:|||:|:
Db   251 VRVSNIEMGPEGATFDIIVSPRAGEAVRYDGLKMPMAGHHNVLNATAAVAVARELGVDAE 310

Qy   284 NIKEALETFGGVKRRFNETTIANQV-IVDDYAHHPREISATIETARKKYPHKEVVAVFQP 342
            |||.||||||||.||||||||:||| ||||||||||| |||||||||||:: :|.:
Db   311 AIAKGLAGFGGVKRPTTTGVANGIRVVDDYGHHPVEIAAVLKAARAVTPNGKVIAVVQP 370

Qy   343 HTFSRTQAFLNEFAEBSLSKADRVFLCEIFGSIRENTGALTIQDLIDKIE-----GASLI 396
            |||||||||||||:|||:||||||| |:|||||||||||||:|||||:::
Db   371 HRFTRLRDIMTEFSSCFNDADTVIVADVYTAGEQPIPGVDRDALVAGLKKFGHRRALPLE 430

Qy   397 NEDSINVL---EQFDNAVILFMGAGDI 420
            ||:|:||. |:||:||||||||||:
Db   431 NPTALPLIAABATSGDIVVLLGGAGDI 457

Search completed: June 3, 2004, 14:41:45
Job time : 21 secs

Qy   378 TGALTIQDLIDKEBGASLIN--------------EDSINVLEQ--FDNAVILFMGAGD 419
            ||||||||||||:.|||||| :||:|||: |||||||||||
Db   410 FNP------DELSVKLFLNIKKINKNTYFFKDVKDSINFIKSLLISGDLFITMGAGN 460

RESULT 14
B81201
UDP-N-acetylmuramate-alanine ligase NMB0423 [imported] - Neisseria meningitidis (strain
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: B81201
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Qin, H.; Vamathevan, J.; Gill, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: B81201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-469 <TET>
A;Cross-references: GB:AE002398; GB:AE002098; NID:g7225640; PIDN:AAF40861.1; PID:g722564
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0423
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match   23.5%;   Score 535;   DB 2;   Length 469;
Best Local Similarity   32.3%;   Pred. No. 3.8e-29;
Matches 150; Conservative 78; Mismatches 194; Indels 42; Gaps 14;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILP-FDANNIK 59
            :|.||||||| |||||:||||||||||:||||||||||||||||||||:|:| :
Db   6 VTNIHFVGIGGVGMSGIAEVLHNLGFKVSGSDQARNAATE-HIGSLGIQVYPGHTAEHVN 64

Qy   60 EDMVVIQGNAFASSHEEIVRAHQLKLDVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLL 119
            : |:||||||||||||||||||||||||:. |||||||||:|:|||||||||||:||||
Db   65 GADVVVTSTAVKENPEVVAALEQQIPVIPRALMLAELMRFRDGIAIAGTHGKTTTTSLT 124

Qy   120 SHVMN-GDKKTSFLIGD-----GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDF 173
            |:. :.|.||||||||. |||:|||||||||||||||:|.|:.:|||:||:|||:
Db   125 ASILGAAGLDPTFVIGGKLNAAGTNARIGKGEYIVAEADESDASFLHLTPIMSVTNIDE 184

Qy   174 DHPD-YFKDINDVFDAFQEMAHNVK--KGIIAWGDDDEHLRKI--EADVPIYYYGFKDSD 227
            ||||.|||||||||||||||||||. ||||: |:.:|:||: |||:|||||||:||
Db   185 DHMDYYGHSVEKLHQAFIDFIHRMPFYGKAFLCI-DSEHVRAILPKVSKPYATYGLDDTA 243

Qy   228 DIYAQNIQITDKGTAFDYV--------DGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLD 281
            |||| |||||.|||||:|. |.:||||||||||||||:||||.|:|:||:
Db   244 DIYATDIENVGAQMKFTVHVQMKGHEQSFEVVLNMP--GRHHNVLNALAAIGVALEVGAS 301

Qy   282 VTNIKEALETFGGVKRRF---NETTIAN--QVIVDDYAHHPREISATIETARKKYPHKE 335
            :|||.||||||||||||: ||||||. |::|||||||||||||||||||:|.:.:
Db   302 VEAIQKGLLGFBGVGRRFQKYGDIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKR 361

Qy   336 VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKI----- 390
            |||||||||||||:::|.||||:||||||:|| |||||||||||||||:||||:|:
Db   362 LVLAFQPHRYTRTRDLFEDFTKVLNTVDALVLTEVYAAGEEPTAAADSRALARAIRVLGK 421

Qy   391 -----EGASLINEDSINVLEQFDNAVILFMGAGDIQKLQNAYL 428
                 |:|||.|||:|||:|.:|||||||||||||||:::|.:
Db   422 LEPIYCENVADLPEMLLNVLQ--DGDIVLNMGAGSINRVPAALL 463

RESULT 15
A87565
UDP-N-acetylmuramate-alanine ligase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87565
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      June  3, 2004, 14:33:06 ; Search time 18 Seconds
                                (without alignments)
                          1264.148 Million cell updates/sec

Title:          US-10-712-713-2
Perfect score:  2275
Sequence:       1 MTHYHFVGIKGSGMSSLAQI..........GDIQKLQNAYLDKLGMKNAF 437

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:      141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2274	100.0	437	1	MURC_STAAM	Q99tc4 staphylococc
2	2255	99.1	437	1	MURC_STAAU	O31211 staphylococc
3	1985	87.3	437	1	MURC_STAEP	Q8c847 staphylococc
4	1470	64.6	436	1	MURC_BACAA	Q8lkq2 bacillus an
5	1470	64.6	436	1	MURC_BACCR	Q816x6 bacillus ce
6	1438	63.2	432	1	MURC_BACSU	P40778 bacillus su
7	1381.5	60.7	438	1	MURC_OCEIH	Q8ep84 oceanobacil
8	1368	60.1	433	1	MURC_BACHD	Q9k7w1 bacillus ha
9	1355	59.6	447	1	MURC_LISIN	Q92ba4 listeria in
10	1342	59.0	447	1	MURC_LISMO	Q8y6s8 listeria mo
11	1284.5	56.5	436	1	MURC_LACPL	Q88wz5 lactobacill
12	1180	51.9	443	1	MURC_LACLA	Q9ce10 lactococcus
13	1139	50.1	442	1	MURC_STRP3	Q8k8j5 streptococc
14	1139	50.1	442	1	MURC_STRPY	Q9a1c7 streptococc
15	1135	49.9	442	1	MURC_STRP8	Q8p2e1 streptococc
16	1119	49.2	444	1	MURC_STRR6	Q8dp40 streptococc
17	1117	49.1	444	1	MURC_STRPN	Q97p88 streptococc
18	701	30.8	460	1	MURC_THETN	Q8r749 thermoanaer
19	641.5	28.2	458	1	MURC_CLOAB	Q97e89 clostridium
20	626.5	27.5	457	1	MURC_CLOPE	Q8xhj0 clostridium
21	596.5	26.2	468	1	MURC_FUSNN	Q8rdq2 fusobacteri
22	566	24.9	495	1	MURC_RICPR	Q92 it9 rickettsia
23	565.5	24.9	485	1	MURC_RICCN	Q92it9 rickettsia
24	548	24.1	469	1	MURC_NEIMA	Q9jsz8 neisseria m
25	538	23.6	468	1	MURC_BORBU	O51757 borrelia bu
26	535	23.5	469	1	MURC_NEIMB	Q9k0y1 neisseria m
27	534	23.5	473	1	MURC_CAUCR	Q9a5a6 caulobacter
28	521.5	22.9	488	1	MURC_SHEON	Q8ep9 shewanella
29	521	22.9	454	1	MURC_AQUAE	O67373 aquifex aeo
30	513	22.5	471	1	MURC_AGRT5	Q8udm9 agrobacteri
31	510.5	22.4	486	1	MURC_CORGL	P94335 corynebacte
32	506	22.2	475	1	MURC_CHLTE	Q8kgd5 chlorobium
33	502	22.1	486	1	MURC_PSESM	Q87wy6 pseudomonas
34	500.5	22.0	494	1	MURC_ANASP	Q8ym75 anabaena sp
35	500	22.0	471	1	MURC_RHIME	Q92mm0 rhizobium m
36	498.5	21.9	475	1	MURC_HAEDU	Q7vmy1 haemophilus
37	496	21.8	802	1	MUDD_CHLMU	Q9plg1 c murc/ddl
38	495.5	21.8	475	1	MURC_WIGBR	Q8d2z7 wiggleswort
39	492.5	21.6	471	1	MURC_BRUME	Q8yi65 brucella me
40	491.5	21.6	475	1	MURC_SYNEL	Q8dlv5 synechococc
41	491.5	21.6	486	1	MURC_BUCBP	P59418 buchnera ap
42	491	21.6	491	1	MURC_ECOL6	Q8x9y7 escherichia
43	491	21.6	491	1	MURC_ECOLI	P17952 escherichia
44	490.5	21.6	491	1	MURC_SALTI	Q8z9g8 salmonella
45	490	21.5	466	1	MURC_RHILO	Q98kb4 rhizobium l

```
                          ALIGNMENTS

RESULT 1
MURC_STAAM
ID   MURC_STAAM     STANDARD;      PRT;    437 AA.
AC   Q99TC4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR SAV1740 OR SA1561 OR MW1683.
OS   Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS   Staphylococcus aureus (strain N315), and
OS   Staphylococcus aureus (strain MW2).
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=158879, 158879, 196620;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mu50 / ATCC 700699, and N315;
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami A., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Kuhara S., Goto S., Yabuzaki J.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001)
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MW2;
RX   MEDLINE=22040717; PubMed=12044378;
RA   Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA   Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA   Yamamoto K., Hiramatsu K.;
RT   "Genome and virulence determinants of high virulence community-
RT   acquired MRSA.";
RL   Lancet 359:1819-1827(2002).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP / UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AP003363; BAB57902.1; -.
DR   EMBL; AP003134; BAB42829.1; -.
```

```
DR   EMBL; AP004828; BAB95548.1; -.
DR   PIR; H89958; H89958.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND       108    114       ATP (POTENTIAL).
SQ   SEQUENCE   437 AA;  49174 MW;  3349943079DB7355 CRC64;

Query Match             100.0%;  Score 2274;  DB 1;  Length 437;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-142;
  Matches  436;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy         61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy        121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy        181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy        241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy        301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy        361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFMGAGDI 420

Qy        421 QKLQNAYLDKLGMKNAF 437
              |||||||||||||||||
Db        421 QKLQNAYLDKLGMKNAF 437

RESULT 2
MURC_STAAU
ID   MURC_STAAU     STANDARD;      PRT;   437 AA.
AC   O31211;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lowe A.M., Deresiewicz R.L.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cell wall formation (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
```

```
     -!- SIMILARITY: Belongs to the murCDEF family.

DR   EMBL; AF034076; AAB87090.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KW   ATP-binding.
FT   NP_BIND       108    114       ATP (POTENTIAL).
SQ   SEQUENCE   437 AA;  49176 MW;  E2FC7D79E6CF8361 CRC64;

Query Match              99.1%;  Score 2255;  DB 1;  Length 437;
  Best Local Similarity   98.9%;  Pred. No. 2.9e-141;
  Matches  432;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60

Qy         61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
              ||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||
Db         61 DMVVIQGNAFASSHEKARAHQMKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120

Qy        121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180

Qy        181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240

Qy        241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300

Qy        301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy        361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||:||||
Db        361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVVLFKGAGDI 420

Qy        421 QKLQNAYLDKLGMKNAF 437
              |||||||||||||||||
Db        421 QKLQNAYLDKLGMKNAF 437

RESULT 3
MURC_STAEP
ID   MURC_STAEP     STANDARD;      PRT;   437 AA.
AC   Q8CS47;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR SE1413.
OS   Staphylococcus epidermidis.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1282;
```

RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 12228;
RX   PubMed=12950922;
RA   Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA   Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA   Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT   "Genome-based analysis of virulence genes in a non-biofilm-forming
RT   Staphylococcus epidermidis strain (ATCC 12228).";
RL   Mol. Microbiol. 49:1577-1593(2003).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC   phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.

DR   EMBL; AE016748; AAO05012.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND   108   114   ATP (POTENTIAL).
SQ   SEQUENCE   437 AA;   49578 MW;   2DB8BFD5D9C1C334 CRC64;

Query Match   87.3%; Score 1985; DB 1; Length 437;
Best Local Similarity   84.7%; Pred. No. 1.7e-123;
Matches 370; Conservative 43; Mismatches 24; Indels 0; Gaps 0;

```
Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  60
         ||||||||||||||||||||||||||||||||::||||||||||||||||||||||: |
Db    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIESVFTEVALRNKGIKILPFDANNITK  60

Qy   61  DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
         :|||||||||  ::||||||||:||||:|::||||:||:||||||||||||||||||||||
Db   61  EMVVIQGNAFPPDNHEEIVRAHELKLDIIKYHDFLGHVINQYTSVAVTGAHGKTSTTGLLS 120

Qy  121  HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
         |||||||||||||||||||||||:|||||||||||||||||:|||||||||||||||||||
Db  121  HVMNGDKKTSFLIGDGTGMGLPGSDYFAFEACEYRRHFLSYHPDYAIMTNIDFDHPDYFK 180

Qy  181  DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQTDKG 240
         :|:||:||||:||| ||||||||||||:|||:: :|:|||||||:.|||||:||||:||
Db  181  NIDDVYDAFQHMALNVKKGIIAWGDDEYLRKLDVDIPVYYYGFKETDDIYAKNIQTEKG 240

Qy  241  TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
         |.|||| .|||||.||||||||.| :||||||||||||||| ::|||||| ||||||||||
Db  241  TQFDVYIKGEFYDQFLSPQYGNHNILNALAVIAISYLEDMNVENIKEALITFGGVKRRFN 300

Qy  301  ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
         ||.: ||||||||||||||||||||||||||| |:.|||||||||||||||||||||||||
Db  301  ETKVSNQVIVDDYAHHPREISATIETARKKYPQKDVVAVFQPHTFSRTQAFLNEFAESLS 360

Qy  361  KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQPDNAVILPMGAGDI 420
         |||:||||||||||||||| |||:||| :|:| | ||:::|:|||: ||||||||||||||
Db  361  KADQVFLCEIFGSIRENTGDLTIEDLINRIDGSTLIDENSIDVLEKFDNAVILPMGAGDI 420

Qy  421  QKLQNAYLDKLGMKNAF 437
         |||::||::|||:||.|
Db  421  QKLLKAYFEKLGVKNDF 437
```

RESULT 4
MURC_BACAA
ID   MURC_BACAA   STANDARD;   PRT;   436 AA.
AC   Q81KQ2;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR BA4938.
OS   Bacillus anthracis (strain Ames).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=198094;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22608414; PubMed=12721629;
RA   Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA   Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA   Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA   Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA   DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA   Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA   Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA   Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA   Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA   Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA   Fraser C.M.;
RT   "The genome sequence of Bacillus anthracis Ames and comparison to
RT   closely related bacteria.";
RL   Nature 423:81-86(2003).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC   phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.

DR   EMBL; AE017039; AAP28622.1; -.
DR   TIGR; BA4938; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND   108   114   ATP (POTENTIAL).
SQ   SEQUENCE   436 AA;   49256 MW;   10C95367937B9FB4F CRC64;

Query Match   64.6%; Score 1470; DB 1; Length 436;
Best Local Similarity   64.7%; Pred. No. 1.3e-89;
Matches 279; Conservative 59; Mismatches 93; Indels 0; Gaps 0;

```
Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  60
         ||.|||||||.|||||||||:||::|||.||| :|: : :||.|:|:||||||:|:||::|
Db    1  MTVYHFVGIKGTGMSSLAQILHDMKGHTVQGSDYEKRFFTQTALEKRNISILPFDKSNVKE  60

Qy   61  DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
         . |:.||||| : :||||.|.:::|:::.|:..|||:::.|||||||||||||||||||||:
Db   61  GQVIAGNAFPDTHEEIVAAKELNIPVHRYHHFLGDLMNQYTSVAVTGAHGKTSTTGLLA 120
```

```
Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||   |||||||||||||| ||| :|||||:|||||||||||||||||| ||||||||||
Db  121 HVMQGAHPTSYLIGDGTGHGVENSKYFVFEACEYRRHFLSYNPDYAIMTNIDFDHPDYFT 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||   ||||| :||||||||:|||| : || :|||||||:|||:|| : |||||||||
Db  181 DINDVFSAFQEMALQVKKGIIACGDDEELQKIQAKVPVIFYGFGEDNDFQARNIQKRTDG 240

Qy  241 TAFDVYDGEFVDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        ||||:|  |||||||||||||:| ||||||||||:| |:|| || : |||||:||||||
Db  241 TIFDVFVRNTYDTFKITGYGNHSVLNALAVIALCHYENVDVEAVKHQLTTFEGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
         ||| : |||||||||||||||||||:||||| : |  : |||||||||| ||:|||||
Db  301 EKPMGEQVIIDDYAHHPTEINATIEAARQKHPEREIVAVFQPHTFSRTEKFLDEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
         ||  |||:||||:||  | ||| || |||: |  |||| :  |: ||||||||||||||
Db  361 KADQVYLCDIFGSARENKGELITEDLQKRIDGAELITDTTTDVLKKHKNGVLIFMGAGDI 420

Qy  421 QKLQNAYLDKL 431
        ||| |||:
Db  421 QKFEAAYVKEV 431
```

```
RESULT 5
MURC_BACCR
ID   MURC_BACCR              STANDARD;       PRT;     436 AA.
AC   Q816X6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR BC4684.
OS   Bacillus cereus (strain ATCC 14579 / DSM 31).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=226900;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22608415; PubMed=12721630;
RA   Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA   Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA   Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
RA   Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA   Overbeek R., Kyrpides N.;
RT   "Genome sequence of Bacillus cereus and comparative analysis with
RT   Bacillus anthracis.";
RL   Nature 423:87-91(2003).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics  and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long  as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement.(See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE017013; AAP11591.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
```

```
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND     108    114       ATP (POTENTIAL).
SQ   SEQUENCE   436 AA;  49200 MW;  70C37FCFBCFFAE74 CRC64;

Query Match             64.6%;  Score 1470;  DB 1;  Length 436;
Best Local Similarity   64.7%;  Pred. No. 1.3e-89;
Matches  279;  Conservative  58;  Mismatches  94;  Indels  0;  Gaps  0;
```

```
Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
        ||||||||||||:||||||||||:|:|||||||||:|:|||||||:|| |: |||||:|||
Db    1 MTVYHFVGIKGTGMSSLAQILHDMKHTVQGSDYEKRFTQTALEKRSISILPFDKNNVEE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
         :||||||||| ||||:|| ||:|:| |  ||||: |  |||||||||||||||||||||
Db   61 GQVIIAGNAFAPDTHEEIVAAKELNIPVRYHHFIGDLMSQYTSVAVTGAHGKTSTTGLLA 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
        |||   |||||||||||||| ||| :|||||:|||||||||||||||||| ||||||||||
Db  121 HVMQGAHPTSYLIGDGTGHGVENSKYFVFEACEYRRHFLSYNPDYAIMTNIDFDHPDYFA 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
        |||||   ||||| :||||||||:|||| : || :|||||||:|||:|| : |||||||||
Db  181 DINDVFSAFQEMALQVKKGIIACGDDEELQKIQAKVPVIFYGFGEDNDFQARNIQKRTDG 240

Qy  241 TAFDVYDGEFVDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
        ||||:|  |||||||||||||:| ||||||||||:| |:|| || : |||||:||||||
Db  241 TIFDVFVRNTYDTFKITGYGNHSVLNALAVIALCHYENVDVEAVKHQLTTFEGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
         ||| : |||||||||||||||||||:||||| : |  : |||||||||| ||:|||||
Db  301 EKPMGEQVIIDDYAHHPTEINATIEAARQKHPEREIVAVFQPHTFSRTEKFLDEFAESLS 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
         ||  |||:||||:||  | ||| || |||: |  |||| :  |: ||||||||||||||
Db  361 KADQVYLCDIFGSARENKGELITEDLQKRIDGAELITDTTTDVLKKHKNGVLIFMGAGDI 420

Qy  421 QKLQNAYLDKL 431
        ||| |||:
Db  421 QKFEAAYVKEV 431
```

```
RESULT 6
MURC_BACSU
ID   MURC_BACSU              STANDARD;       PRT;     432 AA.
AC   P40778;
DT   01-FEB-1995 (Rel. 31, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR BSU29790.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98048467; PubMed=9387221;
RA   Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT   "Sequencing and functional annotation of the Bacillus subtilis genes
RT   in the 200 kb rrnB-dnaB region.";
RL   Microbiology 143:3431-3441(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
```

```
Qy   61 DMVIQGNAFASSHEEIVRAHQLKLDVVSNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
           ||||||  :|| : :|:|:|:|| .:|: |||  :::||||||||:||||||||||:
Db   61 GMTVIAGNAFPDTHPEIEKAMSEGIPVIRYHKFLGDYMKKFTSVAVTGAHGKTSTTGLLA 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          |||: :|||||||||||:|| |:|:|| :||| |||||| |||:|||||||||||||||
Db  121 HVIQNAKPTSFLIGDGTGQGNENSEYFVFEACEYRRHFLSYQPDYAIMTNIDFDHPDYFS 180

Qy  181 DINDVPDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          :|:||||||| ::| ||| |:  || :||||||:|:||:|| ||  : : ::|
Db  181 SIDDVPDAFQEMALQVNKGIIACGDDEHLPKIHANVPVYYGTGEENDFQARNIVKSTEG 240

Qy  241 TAFDVYVDGBFYDHFLSPQYGDHTVLNALAVTAISYLEKLDVTNIKEALETFGGVKRRFN 300
          |||: |  |||||||:|::||||||||||||||:  ||||||||:  : :|:| |:||||
Db  241 TTFDVFVRNTFYDTFYIPAYGHHNVLNSLAVIALCHYEBIDSSIIKHALKSFGGVKRRFN 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRIQAFLNEFAESLS 360
          |  ::| ||:|||||:|| :|:|  ||  |  :||::||||||||::| |||:|:||||||
Db  301 EKQLGDQVLIDDYAHHPTBIKVTIEAARQKYPDREIVAVFQPHTFRTQQFLDEFAESLS 360

Qy  361 KADRVPLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          ||  ||| :|||| |||  |||:|||:|:||||| :|:|||  ||:| :||:|||||||:
Db  361 GADCVYLCDIFGSARRNAGKLTIGDLQGKIHNAKLIEBDDTSVLKAHDKAVLIFMGAGDI 420

Qy  421 QKLQNAY 427
          ||:| ||
Db  421 QKYMRAY 427
```

```
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [3]
RP   SEQUENCE OF 85-432 FROM N.A.
RC   STRAIN=168 / Marburg;
RX   MEDLINE=96310371; PubMed=8733232;
RA   Varon D., Brody M.S., Price C.W.;
RT   "Bacillus subtilis operon under the dual control of the general
RT   stress transcription factor sigma B and the sporulation transcription
RT   factor sigma H.";
RL   Mol. Microbiol. 20:339-350(1996).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
CC
```
```
DR   EMBL; AF008220; AAC00294.1; -.
DR   EMBL; Z99119; CAB14957.1; -.
DR   EMBL; L31845; AAB40043.1; -.
DR   PIR; C69662; C69662.
DR   SubtiList; BG10973; murC.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KW   ATP-binding; Complete proteome;
FT   NP_BIND  108  114  ATP (POTENTIAL).
SQ   SEQUENCE  432 AA; 48364 MW; B73BF9502FD7CE1B CRC64;
```

```
RESULT 7
MURC_OCEIH
ID   MURC_OCEIH    STANDARD;    PRT;   438 AA.
AC   Q8EP84;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR OB2231.
OS   Oceanobacillus iheyensis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX   NCBI_TaxID=182710;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HTB831 / DSM 14371 / JCM 11309;
RX   MEDLINE=22220767; PubMed=12235376;
RA   Takami H., Takaki Y., Uchiyama I.;
RT   "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT   Ridge and its unexpected adaptive capabilities to extreme
RT   environments.";
RL   Nucleic Acids Res. 30:3927-3935(2002).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
CC
```
```
DR   EMBL; AP004600; BAC14187.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
```

```
Query Match     63.2%;  Score 1438;  DB 1;  Length 432;
Best Local Similarity   64.6%;  Pred. No. 1.6e-87;
Matches  276;  Conservative  51;  Mismatches 100;  Indels   0;  Gaps   0;

Qy   1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
         ||||||||:|||||:|||:|| || ||:||||||  || :| |:||||| |:: |:|||
Db   1 MTVYHFVGIKGTGMSSLAQILHDNGYTVQGSDIEKFIFTQTALEKRNITILPFSAENIKP 60
```

```
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND      108    114       ATP (POTENTIAL).
SQ   SEQUENCE   438 AA;  49863 MW;  43AE610786AB1992 CRC64;

Query Match             60.7%;  Score 1381.5;  DB 1;  Length 438;
Best Local Similarity   61.1%;  Pred. No. 8.7e-84;  Indels 1;  Gaps 1;
         Matches 261;  Conservative 63;  Mismatches 102;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          |||||||||||||||||:||:||||||||:|||: :||||||||||||||||||||||||
Db    1 MTTYHFIGIKGTGMSALAQILHDSGEKVQGSDFEKRFFTQBALEQKNITILPFSKENIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
          ::| ||||||:   |||| :    |:|:|  ||||||:|||||:|||||||||||||||||
Db   61 DYTIIAGNAFSDDHIEIQKAKELGCKFYRHEFIGEWLKQYTSIAVTGAHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||:.  : ||||:|||||||: ::|||:|||||||||||.| :||||||||||||||||:
Db  121 HVLKSAYPISYILGDGTGNGHVDSEYFVFEACEYRRHFLKYEPDYAIMTNIDFDHPDYFT 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          .:.||.||||.||..|||.|:|.||||::|.| : ||:|.:||:||.|::|:|: : ..|
Db  181 SLDDVVDAFQSMADRVKKCIVACGDDEQLQGIHTKVPVIYGFNDSNDFQAQNVVETANG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          |  ||:|   .|||:: | ||:|||||||||||:|:||   |::::|:  |:|||||||:
Db  241 TEPDVFVRNTYDHFIIPTYGNHTVLNALAVIAICHYEGISVEHMKQ-LDTFEGVKRRFT 299

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          | .: . :|:|||||||:|||.||| |||||||::.||.|||||| ||: ||.|||:||:
Db  300 EKEIGQSIVIDDYAHHPKEISVTIESARKKYPNQNVVAIFQPHTFTRTKTFLKEFADSLN 359

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          :|| |:||:||.| ||:.| |.|.||::|| .:.|: :   .:|:.|::.|::||||||||
Db  360 EADNVYLCDIFKSAREDSGQLITNDLQKLIPNSQLLSLKETEVLQBFKDSVLIFMGAGDI 419

Qy  421 QKLQNAY 427
         |||:.|| |
Db  420 QKFQAAY 426

RESULT 8
MURC_BACHD
ID   MURC_BACHD     STANDARD;      PRT;      433 AA.
AC   Q9K7W1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR BH3248.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------
DR   EMBL; AP001518; BAB06967.1; -.
DR   PIR; H84055; H84055.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND      108    114       ATP (POTENTIAL).
SQ   SEQUENCE   433 AA;  48656 MW;  A5E6962FCA4760B7 CRC64;

Query Match             60.1%;  Score 1368;  DB 1;  Length 433;
Best Local Similarity   59.3%;  Pred. No. 6.6e-83;  Indels 0;  Gaps 0;
         Matches 253;  Conservative 69;  Mismatches 105;

Qy    1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
          ||:.||:||||||||:|||:|||:.: :||||||:|:||.:| ::||:|:|.||:|:|||
Db    1 MTTFHFIGIKGSGMSALAQVLHDMNYQVQGSDVEKRFFTQKPLEEKGIPIFPFQQKENIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
           .  |  :|:   |||: .|::  | :  || ||| |  | :| |||||:||||||||||||
Db   61 GQQVVVSAAYSDHEEVVAAKEQGLEIQSYPTFLGEFIQRFTSIAVTGSHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAPEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
          ||:    .||||||||||.|.. |.|||:||||||||||.|:||| ||||||||||||||
Db  121 HVLGSHCPTSFLIGDGTGKGAADSHYFVFEACEYRRHFLNYRPDYCIMTNIDFDHPDYFD 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
          ||||||.||||||..|||.|:|.||||:|:.| |.||::| ||::.|.: | | :.:.:|
Db  181 DINDVFSAFQEMAMQVKKAIVACGDDEQLQGIHANVPVVYGFNEDNDFQARNIEVTSEG 240

Qy  241 TAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
          | ..|:|   .||||::| ||:||||:||||||:|:||:: .|:: |||.|:||||||||
Db  241 THFQVFVRNNHYGDFTIPRFGKHNVLNSLAVIALCHYENIPVEVLARELKTFEGVKRRFT 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
          |  : .|:::|||||||.|||||.| ||:|:||.:.||.|||||||||:|||.|||:.|:
Db  301 EKTLGGQILIDDYAHHPTEISATTEAAAKQKYPDRPVVAIFQPHTFSRTKAFLHEFAEALK 360

Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
          :||.|:||:||||:||:.|:|:|:||: .| .|.|::|:.  .||: ::: |:|||||||
Db  361 QADHVYLCDIFGSARENKGSLSIEDLQALIPGAHLLSENDLAPLQDHRSGVLLFMGAGDI 420

Qy  421 QKLQNAY 427
         |||:.|| |
Db  421 QKFQQAY 427

RESULT 9
MURC_LISIN
ID   MURC_LISIN     STANDARD;      PRT;      447 AA.
AC   Q92BA4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR LIN1646.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
```

```
Qy  361 KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
            ||||| ||::|: |:|||  || ||:||:|  :|| | ||||||||||||
Db  361 LADEVVLCDIFGSAREKTGNLTIADLAHKTKGNHIIKEEHTELLQYPEAVILFMGAGDV 420

Qy  421 QKLQNAY 427
            || ||||
Db  421 QKFQAAY 427
```

RESULT 10
MURC_LISMO
ID   MURC_LISMO   STANDARD;   PRT;   447 AA.
AC   Q8Y6S8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR LMO1605.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-e / Serovar 1/2a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
-!- FUNCTION: Cell wall formation.
-!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
-!-   phosphate + UDP-N-acetylmuramoyl-L-alanine.
-!- PATHWAY: Peptidoglycan biosynthesis.
-!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
-!- SIMILARITY: Belongs to the murCDEF family.
```

```
DR   EMBL; AL591979; CAC99683.1; -.
DR   PIR; AE1275; AE1275.
DR   ListList; LMO1605; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND     108    114       ATP (POTENTIAL).
SQ   SEQUENCE   447 AA;  50073 MW;  37BA8AFB45F31994 CRC64;

Query Match             59.0%;  Score 1342;  DB 1;  Length 447;
Best Local Similarity   59.7%;  Pred. No. 3.6e-81;
Matches 255; Conservative  64; Mismatches 108; Indels  0; Gaps  0;
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / Serovar 6a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
-!- FUNCTION: Cell wall formation.
-!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
-!-   phosphate + UDP-N-acetylmuramoyl-L-alanine.
-!- PATHWAY: Peptidoglycan biosynthesis.
-!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
-!- SIMILARITY: Belongs to the murCDEF family.
```

```
DR   EMBL; AL596169; CAC96877.1; -.
DR   PIR; AE1638; AE1638.
DR   ListList; LIN01646; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND     108    114       ATP (POTENTIAL).
SQ   SEQUENCE   447 AA;  50036 MW;  AFA793A96833E904 CRC64;

Query Match             59.6%;  Score 1355;  DB 1;  Length 447;
Best Local Similarity   60.2%;  Pred. No. 5e-82;
Matches 257; Conservative  63; Mismatches 107; Indels  0; Gaps  0;
```

```
Qy    1 MTHYHFVGIKGSGMSSLAQIMEDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE 60
           |||||||||||||||  |:|| || |:|||::  |:: ||  |:|:||| |:|:|||
Db    1 MTIYHFVGIKGSGMSALAQILHDKGFQVQGSDVDKYFFTQKALEEKQIPIMTFSADNIKE 60

Qy   61 DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
           :||||| :  :: ||| |: |||: || |: | | ||:::|:  :||| |||||||||
Db   61 GLTIIAGNAFPDTHEEIERANELNLPVIRYHKFPLGQLIDGYTSIAITGSHGKTSTTGLLS 120

Qy  121 HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
           ||: ::| ||| |||:| ||| |:|| | ||| |:|||  :||||:|||:|:|||||||
Db  121 HVVGAIRPTSYLIGDGTGKDAKYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK 180

Qy  181 DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
           :||  || :|:  |  || ||:| |||: ||:: ||||:|  :| || || :|: ||
Db  181 SVDDVFNAFETIGKQVKKAVFALGDDVELRKLSLDVPIIYFGFGSENEFQAKNVKKEYTG 240

Qy  241 TAFDVVVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRFPN 300
           ||:||||| : ::|| | || |:||| ||:: ||::  ::| |:|  ::| ||||||
Db  241 TKFDVVHRDEFLASFEIPAYGDHNVLNALSVIALCDYEGLPVDEVKKELKTFPGVKRFS 300

Qy  301 ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLS 360
           ||: |||:|:|||||| |:||:||: ||:||:|: |||||||||:|||: || |::||
Db  301 ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPNKKIVAVFQPHTFTRTRTFLQGFADSLN 360
```

```
CC
DR   EMBL; AL935256; CAD63921.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND   111   117   ATP (POTENTIAL).
SQ   SEQUENCE   436 AA;   48750 MW;   25A582D850238CD4 CRC64;

Query Match             56.5%;   Score 1284.5;   DB 1;   Length 436;
Best Local Similarity   55.7%;   Pred. No. 2.1e-77;
Matches  240;   Conservative   75;   Mismatches 115;   Indels   1;   Gaps   1;

Qy    2  THYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKED  61
            |:||||||||||||||| :| :|:||| ||||||:||: ||:: ||  :|||||||
Db    5  TVYYFVGIKGSGMSSLALILHDKGYVEGSDIEQYTFTQKGLAAAGIKMLPFSEDNIREG  64

Qy   62  MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
            : ||| :|:  ||| ::  |||:  :::||| |: ||:|:|| :|:|||||||||||||
Db   65  LTVIAGNSFTDDHPEIKKAREMGLPVRYRHEFLGKLMEGFTSIGVAGTHGKTSTTGLLSH 124

Qy  122  VMNGDKKTSFLIGDGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDHPDYFKD 181
            |: | ||:|:|||||| |:  |::| : ||||||||:::  ||||||||:|||||||:||
Db  125  VLSHIAPTSYLIGDGTGKGTPDARFFVFEADEYRRHFVAYHPDYAIMTNVDFDHPDYYKD 184

Qy  182  INDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGT 241
            ::|:|:||:: :|:||||:|||||||:|::::|| |||||| :|:||  | :|::  |||
Db  185  LADVQSAFQQFGNQVKKGIFAWGDDESLRHLDVDTPIYYYGTNDRDDFQAVNIKRTTKGS 244

Qy  242  AFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN  301
            :|:|  |  : : | |   |:|:||:  ||||:||:||:::: ||: ||  |:|||||::
Db  245  SFEVKYHDESLGKFBIPLFGEHNVLNSTAVIAVSYFEKVNLDEIRRELLDFSGVKRRFSE 304

Qy  302  TTIANQVIVDDYAHHPREISATTETARKKYPHKEVAVFQPHTFSRTQAFLNEFAESLSK  361
            : : : ||:||||||| ||:|| :|:::|||::| |:|||||||||:|:: :|||:|||
Db  305  HQVGDMVMIDDYAHHPSEIKATLDAARQKYPDKEILAVFQPHTFSRTKALMDGFAASLSK 364

Qy  362  ADRVFLCBIFGSIRENTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDI 420
            || |||  || | :|:  :::::| :::|::  ::  || :      :|:|:||||||||
Db  365  ADHVFLTNIFSSAREKSGDVSSKDLAAKLPNGGEIITTDDMSALTAYHNAVAVFMGAGDI 424

Qy  421  QKLQNAYLDKL 431
            ||  :: :
Db  425  QKYEKIYEDQM 435


RESULT 12
MURC_LACLA
ID   MURC_LACLA   STANDARD;   PRT;   443 AA.
AC   Q9CE10;
DT   16-OCT-2001 (Rel. 40, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR LL2040.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
```

```
Qy    1  MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKE  60
            ||:||||||||||||||||  :| :|:||| ||||||:||: ||:: ||  :|||||||
Db    1  MTIYHFVGIKGSGMSSLALAQILHDKGFQVQGSDVDKYFFTQKALEEKQIPIMTFSADNIQE  60

Qy   61  DMVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS 120
            | : ||| :|   ||| ::  :|:  :::||| |: |:|:|||:|:|:|||||||||||
Db   61  GLTIIAGNAFPDTHEEIERALEGLSVIRYHKFLGQLIDGYTSIAITGSHGKTSTTGLLS  120

Qy  121  HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFK 180
            || : | ||:|:|||||| |:  |:::|: ||||:|||:::  ||||||||:|||||||
Db  121  HVGAIRPTSYLIGDGTGSGTKDAEYFALEACEYQRHFLAYKPTYAIMTNIDWDHPDYFK  180

Qy  181  DINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKG 240
            ::::||:|| : :|:||| :|||||| |:|::: || |:|:|||     : :|:: : :||
Db  181  SVDDVFNAFETLGKQVKKAVFALGDDAELRKLTLDIPIIYFGFGEENERQAKRNVIKETTG 240

Qy  241  TAPDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFN 300
            | ||||  : ||  | |  ||:| :||||:||| | ||  |  ::::| ||  |:|||||
Db  241  TKFPDVYHREEFLSSFEIPAYGDHNVLNALSVIALCDYEGLPVEDVKNELKTFEGVKRRFS 300

Qy  301  ETTIANQVIVDDYAHHPREISATIETARKKYPHKEVAVFQPHTFSRTQAFLNEFAESLS  360
            : |  ||||:||||||| ||:|  :|:::|||:: :|:||||||:||:::::| |:||:
Db  301  ITEKGNQVLVDDYAHHPSEIRATVNAARQKYPDKKVVAVFQPHTFTRTRTFLQGFADSLN 360

Qy  361  KADRVFLCEIFGSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFDNAVILFMGAGDI 420
            :||:|| |:|||| :|:| | ::  |::: |:  ::::|:::::   ::|||||||||||:
Db  361  LADEVYLCDIFGSGAREKTGNLTIADLAHKTKGNHIIKEHTEELLKYPEAVILFMGAGDV 420

Qy  421  QKLQNAY 427
            ||:| ||
Db  421  QKFQAAY 427


RESULT 11
MURC_LACPL
ID   MURC_LACPL   STANDARD;   PRT;   436 AA.
AC   Q88W25;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR LP_1462.
OS   Lactobacillus plantarum.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1590;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIMB 8826 / WCFS1;
RX   MEDLINE=22480296; PubMed=12566566;
RA   Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA   Kuipers O.P., Leer R.J., Tarchini R., Peters S.A., Sandbrink H.M.,
RA   Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA   Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA   De Vos W.M., Siezen R.J.;
RT   "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
CC      phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
DR   EMBL; AE006433; AAK06138.1; ALT_INIT.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND     110    116    ATP (POTENTIAL).
SQ   SEQUENCE   443 AA;   50050 MW;   ABEDAB52E9638C8A CRC64;

Query Match      51.9%;  Score 1180;  DB 1;  Length 443;
Best Local Similarity   52.7%;  Pred. No. 1.6e-70;
Matches  227;  Conservative  79;  Mismatches 121;  Indels 4;  Gaps 4;

Qy    4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV   63
            ||| ||||||||| :|||  :| |||||| :||: || : :||||  :||| |  :||||
Db    5  YHFTGIKGSGMSALALMLHQMGHKVQGSDSTDYFFTQRGLEQADVPLLPFDEKNIKPEFE   64

Qy   64  VIQGNAFASSHE-EIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV  122
Db   65  LIAGNAFRDDNNVETAFAHKNGFPFKRYHEFLGHFMEDFTSIGVAGAHGKTSTTGMLAHV  124

Qy  123  MNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI  182
Db  125  MSNIVDTSYLIGDGTGRGIEGSEYFVFESDEYERHFMPYHPEYTIMTNIDFDHPDYFEGI  184

Qy  183  NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTA  242
Db  185  EDVTSAFQDYANNIKKGIFAYGEDVNLRKLTAKAPIIYYGFEANDDYRAENLVRSTRGSS  244

Qy  243  FDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKRALETFGGVKRRFNET  302
Db  245  FDAYFPRGEKIGHFVVPAYGKHVLNALSVVAVCHNLGLDMTEVADHLLTFRGVKRRFTEK  304

Qy  303  TIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAPLNEFAESLSKA  362
Db  305  KVGETVIIDDFAHHPTEIEATLDAARQKYPDREIVAVFQPHTFRTIAFADEFAEVLDHA   364

Qy  363  DRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVLFMGAGDI  420
Db  365  DTVYLAQIYGSAREVDHHEITAQDLADKVRKPAKVIDLDNVSPLLDHDRGVVVFMGAGNI  424

Qy  421  QKLQNAYLDKL  431
Db  425  QKYELAF-EKL  434

RESULT 13
MURC_STRP3
ID   MURC_STRP3   STANDARD;   PRT;   442 AA.
AC   Q8K8J5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR SPYM3_0252 OR SPS1607.
```

```
OS   Streptococcus pyogenes (serotype M3).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=198466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS315 / Serotype M3;
RX   MEDLINE=22133808; PubMed=12212206;
RA   Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
RA   Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA   Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA   Schlievert P.M., Musser J.M.;
RT   "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT   phage-encoded toxins, the high-virulence phenotype, and clone
RT   emergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSI-1 / Serotype M3;
RX   MEDLINE=22683278; PubMed=12799345;
RA   Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA   Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA   Hayashi H., Hattori M., Hamada S.;
RT   "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT   large-scale genomic rearrangement in invasive strains and new insights
RT   into phage evolution.";
RL   Genome Res. 13:1042-1055(2003).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
DR   EMBL; AE014141; AAM78859.1; -.
DR   EMBL; AP005146; BAC64702.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND    109    115    ATP (POTENTIAL).
SQ   SEQUENCE   442 AA;   49622 MW;   C0216CAEFF922A03 CRC64;

Query Match      50.1%;  Score 1139;  DB 1;  Length 442;
Best Local Similarity   50.7%;  Pred. No. 7.9e-68;
Matches  218;  Conservative  82;  Mismatches 128;  Indels 2;  Gaps 2;

Qy    4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV   63
            ||| ||||||||| :|||  :| |||||| :||:  || :   :||| |  :||||
Db    5  YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYYFTQRGLEQAGITILPFSEDNITPDME   64

Qy   64  VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM  123
Db   65  LIVGNAFRENNKEVAYALRHQIPFKRYHDFLGDFMKSPISPAVAGAHGKTSTTGLLSHVL  124

Qy  124  NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN  183
Db  125  KNITDTSYLIGDGTGRGSANAQYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTGIA  184

Qy  184  DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF  243
```

Left column (RESULT 14)

```
                   ||:||   |||    :|: .||:|| ||||||::: :| :||||||  |::|
Db  185 DVRNAFNDYAKQVKKALFVYGEDDELKKIEAPAPIYYYGFEEGNDFIAYDITRTTNGSDF 244

Qy  244 DYYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
            :| :||||||  ::   ||||  .:|: . ||| :: |.:||||||||
Db  245 KVKHQGEVIGQFHVPAYGKHNILNATAVIANLFVAGIDMALVADHLKTFSGVKRRFTEKI 304

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
         : :||||| ||||.|| || :|.|:|||| |:|||||||| .: .|:: |.:| ::||
Db  305 INDTIIIDDFAHHPTEIVATIDAARQKYPSKEIVAIFQPHTFRTIALLEDFACAlNEAD 364

Qy  364 RVFLCEIFPGSIRE-NTGALTIQDLLIDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQ 421
          |:| :|: :|:|  ::  : : :|.  ::::         :::     .||:|.||||||||
Db  365 SVYLAQIYGSAREVDKGEVKVEDLAAKIIKPSQVVTVENVSPLLDHDNAVYVFMGAGDIQ 424

Qy  422 KLQNAYLDKL 431
         :::: :::|
Db  425 LYEHSFEELL 434
```

```
RESULT 14
MURC_STRPY
ID   MURC_STRPY     STANDARD;      PRT;    442 AA.
AC   Q9A1C7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR SPY0345.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
DR   EMBL; AE006499; AAK33396.1; -.
DR   HAMAP; MF_00046; -; 1.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR005758; MurC.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; ATP-binding; Cell division; Cell wall;
KW   Peptidoglycan synthesis; Complete proteome.
FT   NP_BIND     109    115    ATP (POTENTIAL).
SQ   SEQUENCE   442 AA;  49594 MW;  C02175B7F8B3303 CRC64;

Query Match             50.1%;  Score 1139;  DB 1;  Length 442;
```

Right column (RESULT 15)

```
Best Local Similarity  50.7%;  Pred. No. 7.9e-68;
Matches 218;  Conservative 82;  Mismatches 128;  Indels 2;  Gaps 2;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKILPFDANNIKEDMV 63
         |||.|||||||| :| :| :.|||:|||||:|: :||.|.  |:||||| : ::| | ||
Db    5 YHFIGIKGSGMSALALMLHQMGHKVQGSDVEKYYFTQRCLEQAGITILPFSEDNITPDME 64

Qy   64 VIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGGLLSHVM 123
         :: ||||  .::||.:.|: | |  ||.|||. .:.|  |||||||||||||||.||||:
Db   65 LIVGNAFRENNKEVAYALRHQIPFKRYHDFLGDFMKSFISFAVAGAHGKTSTTGLLSHVL 124

Qy  124 NGDKKTSFLIGDTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
          .  :|.|||||||.| :.::| .|:: :|||||: |.||::||:||||||||||| .|
Db  125 KNITDTSYLIGDTGRGSANAQYFVFESDEYERHFMPYVHPEYSIITNIDFDHPDYFTGIA 184

Qy  184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
         || :||::   |||| | :|.||| |:||||  |||||||::.:| || :: | |.|  |
Db  185 DVRNAFNDYAKQVKKALFVYGEDDELKKIEAPAPIYYYGFEEGNDFIAYDITRTTNGSDF 244

Qy  244 DYYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
            :| :||||||  ::   ||||  .:|: . ||| :: |.:||||||||
Db  245 KVKHQGEVIGQFHVPAYGKHNILNATAVIANLFVAGIDMALVADHLKTFSGVKRRFTEKI 304

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
         : :||||| ||||.|| || :|.|:|||| |:|||||||| .: .|:: |.:| ::||
Db  305 INDTIIIDDFAHHPTEIVATIDAARQKYPSKEIVAIFQPHTFRTIALLEDFACAlNEAD 364

Qy  364 RVFLCEIFPGSIRE-NTGALTIQDLLIDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQ 421
          |:| :|: :|:|  ::  : : :|.  ::::         :::     .||:|.||||||||
Db  365 SVYLAQIYGSAREVDKGEVKVEDLAAKIIKPSQVVTVENVSPLLDHDNAVYVFMGAGDIQ 424

Qy  422 KLQNAYLDKL 431
         :::: :::|
Db  425 LYEHSFEELL 434
```

```
RESULT 15
MURC_STRP8
ID   MURC_STRP8     STANDARD;      PRT;    442 AA.
AC   Q8P2E1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8) (UDP-N-
DE   acetylmuramoyl-L-alanine synthetase).
GN   MURC OR SPYM18_0398.
OS   Streptococcus pyogenes (serotype M18).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=186103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS8232 / Serotype M18;
RX   MEDLINE=21927593; PubMed=11917108;
RA   Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA   Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA   Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA   Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT   "Genome sequence and comparative microarray analysis of serotype M18
RT   group A Streptococcus strains associated with acute rheumatic fever
RT   outbreaks.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC   -!- FUNCTION: Cell wall formation.
CC   -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +
CC       phosphate + UDP-N-acetylmuramoyl-L-alanine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the murCDEF family.
```

```
CC    modified and this statement is not removed.  Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------------
DR    EMBL; AE009982; AAL97144.1; -.
DR    HAMAP; MF_00046; -; 1.
DR    InterPro; IPR000713; Mur_ligase.
DR    InterPro; IPR004101; Mur_ligase_C.
DR    InterPro; IPR005758; MurC.
DR    Pfam; PF01225; Mur_ligase; 1.
DR    Pfam; PF02875; Mur_ligase_C; 1.
DR    TIGRFAMs; TIGR01082; murC; 1.
KW    Ligase; ATP-binding; Cell division; Cell wall;
KW    Peptidoglycan synthesis; Complete proteome.
FT    NP_BIND      109    115       ATP (POTENTIAL).
SQ    SEQUENCE   442 AA;  49583 MW;  52535CCDDA3356B7 CRC64;
```

```
Query Match          49.9%;  Score 1135;  DB 1;  Length 442;
Best Local Similarity   50.5%;  Pred. No. 1.5e-67;
Matches  217;  Conservative   82;  Mismatches  129;  Indels    2;  Gaps    2;

Qy      4  YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV    63
           |||:||||||||| ||  :|| ||||||:|:| :|| ||| :|| :||
Db      5  YHFIGIKGSGMSALALMLHQMGHMVQGSDVEKYYFTQRGLEQAGITILPFSEDNITPDME    64

Qy     64  VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM   123
           :| |||| |  : ::     |:||||||||||||||||||||||
Db     65  LIVGNAFRENNKEVAYALRHQIPFKRYHDFLGDFMKSFISFAVAGAHGKTSTTGLLSHVL   124

Qy    124  NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN   183
           |||  ||||||| | |||  :|| :|||||:| :|:|:||||||||||||||||||
Db    125  KNITDTSYLIGDGTGRGSANAQYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTGIA   184

Qy    184  DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYGFKDSDDIYAQNIQITDKGTAF   243
           ||  ||:|  ||     ||| ||  :|||||  |:||| :||:||
Db    185  DVRNAFNDYAKQVKKALFVYGEDDELKKIEAPAPIYYYGFEEGNDFIAYDITRTTNGSDF   244

Qy    244  DVYYDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT   303
           ||||||| |||  :   ||:  :||:|||:| |:| ||||||
Db    245  KVKHQGEVIGQFHVPAYGKNILNATAVIANLFVAGIDMALVADHLKTFSGVKRRFTEKI   304

Qy    304  IANQVIVDDYAHHPRISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD   363
           : :|||||||||:|:| ||||||| ||||| ||||:||||| :|::|| ::|:|:|
Db    305  VNDYIIDDFAHHPTRIVATIDAARQKYPSKEIVAIFQPHTFTRTIALLEDFACALNEAD   364

Qy    364  RVFLCEIFGSIRE-NTGALTIQDLIDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQ   421
           |:| :||| |:| |     | :|||      : :::  | ||||||||||||||||||
Db    365  SVYLAQIYGSAREVDKGEVKVEDLAAKIIKPSQVVTVENVSPLLDHDNAVYVFMGAGDIQ   424

Qy    422  KLQNAYLDKL   431
           :|::|:|
Db    425  LYEHSPEELL   434
```

Search completed: June 3, 2004, 14:39:38
Job time : 19 secs

```
              GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    June  3, 2004, 14:36:31 ; Search time 46 Seconds
                              (without alignments)
                              2997.423 Million cell updates/sec

Title:           US-10-712-713-2
Perfect score:   2275
Sequence:        1 MTHYHFVGIKGSGMSSLAQI.........GDIQKLQNAYLDKLGMKNAF 437

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:     1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_25:*
              1: sp_archea:*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
             15: sp_rvirus:*
             16: sp_bacteriap:*
             17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1438.5	63.2	445	16	Q833N6	Q833n6 enterococcu
2	1284.5	56.5	436	16	Q88WZ5	Q88wz5 lactobacill
3	1163.5	51.1	443	16	Q8DY77	Q8dy77 streptococc
4	1153.5	50.7	443	16	Q8E3U2	Q8e3u2 streptococc
5	1132.5	49.8	452	16	Q8DSP4	Q8dsp4 streptococc
6	633.5	27.8	456	16	Q899J1	Q899j1 clostridium
7	536.5	23.6	486	2	Q9RNM7	Q9rnm7 zymomonas m
8	523.5	23.0	494	2	Q9FlM9	Q9flm9 shewanella
9	521.5	22.9	488	16	Q8E9P8	Q8e9p8 shewanella
10	506.5	22.3	468	16	Q7WFS3	Q7wfs3 bordetella
11	506.5	22.3	468	16	Q7W4B5	Q7w4b5 bordetella
12	506.5	22.3	468	16	Q7VUQ4	Q7vuq4 bordetella
13	502	22.1	486	16	Q87WY6	Q87wy6 pseudomonas
14	500	22.0	465	16	Q83F17	Q83f17 coxiella bu
15	500	22.0	473	16	Q82VS2	Q82vs2 nitrosomona
16	499.5	22.0	473	16	Q7V3P8	Q7v3p8 prochloroco
17	498.5	21.9	472	16	Q7VEJ1	Q7vej1 prochloroco
18	498.5	21.9	475	16	Q7VMY1	Q7vmy1 haemophilus
19	490	21.5	487	2	Q9RGR6	Q9rgr6 pseudomonas
20	490	21.5	491	16	Q83MF8	Q83mf8 shigella fl
21	488.5	21.5	482	16	Q88N75	Q88n75 pseudomonas
22	484.5	21.3	512	16	Q8G4Q4	Q8g4q4 bifidobacte
23	480	21.1	464	16	Q82I07	Q82107 streptomyce
24	474.5	20.9	467	16	Q89FU8	Q89fu8 bradyrhizob
25	464	20.4	811	16	Q82I84	Q82i84 chlamydophi
26	462.5	20.3	190	2	Q9KGV8	Q9kgv8 lactococcus
27	462	20.3	467	16	Q8A259	Q8a259 bacteroides
28	456	20.0	488	16	Q7V9C3	Q7v9c3 prochloroco
29	452.5	19.9	467	16	Q7UA71	Q7ua71 synechococc
30	449	19.7	474	16	Q8KCE3	Q8kce3 chlorobium
31	439.5	19.3	457	16	Q9CMA5	Q9cma5 pasteurella
32	426	18.7	439	16	Q7VJ70	Q7vj70 helicobacte
33	424	18.6	459	16	Q7VN73	Q7vn73 helicobacte
34	420.5	18.5	461	16	Q8XVP8	Q8xvp8 ralstonia s
35	412.5	18.1	480	16	Q7VQI7	Q7vqi7 candidatus
36	405	17.8	484	16	Q8DE91	Q8de91 vibrio vuln
37	399	17.5	452	16	Q8PH24	Q8ph24 xanthomonas
38	398	17.5	452	16	Q8P5P6	Q8p5p6 xanthomonas
39	393.5	17.3	475	16	Q83GN1	Q83gn1 tropheryma
40	393.5	17.3	490	16	Q8F4J0	Q8f4j0 leptospira
41	392.5	17.3	475	16	Q83HK2	Q83hk2 tropheryma
42	392	17.2	452	16	Q87SV8	Q87sv8 vibrio para
43	390	17.1	458	16	Q9JUE5	Q9jue5 neisseria m
44	383	16.8	458	16	Q9JRY9	Q9jry9 neisseria m
45	382	16.8	465	16	Q8EYR2	Q8eyr2 leptospira

ALIGNMENTS

```
RESULT 1
Q833N6
ID   Q833N6           PRELIMINARY;    PRT;   445 AA.
AC   Q833N6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase.
GN   MURC OR EF1908.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
DR   EMBL; AE016953; AAO81660.1; -.
DR   TIGR; EF1908; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
```

KW   Ligase; Complete proteome.
SQ   SEQUENCE   445 AA; 49999 MW; 12189004F4FB1E9A CRC64;

```
Query Match             63.2%;
Best Local Similarity   60.9%;    Score 1438.5;  DB 16;  Length 445;
Matches 259; Conservative 79; Mismatches 86;  Indels 1; Pred. No. 1.2e-92;  Gaps 1;

Qy    4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
          |||||||||||||:| :||:|| ||:|||||:| :||||  :|::| |:| |:|:|||:
Db    9 YHFVGIKGSGMSSLALVLHQKGYNVQGSDVEYFFTQRDLEKSGVPILPFNADNIDKDMI 68

Qy   64 VIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
         ||:|||| :|||||  |:| | :|: :|:||   |||||:|||:||||||||||||:|| 
Db   69 VIAGNAFPDTHEEIARAIELGAEVIRYHDFIARFIEPYTSIAVTGSHGKTSTTGLLAHVL 128

Qy  124 NGDKKTSFLIGDGTGMGLPESDYPAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
          |:  TS:LIGDGTG G P:|  :AFEACEYRRHFL:Y PDYAIMTNIDFDHPDY:K:I:
Db  129 SGINPTSYLIGDGTGHGEPDADFFAFEACEYRRHFLAYSPDYAIMTNIDFDHPDYYKSIE 188

Qy  184 DVFDAFQEMARHVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGTAF 243
         |||.|| :|| :|||||:|:|||::||::|::VP:YYYG :::DDI|| NIQ|:::.::F
Db  189 DVFSARFQTMAHQVKKGIFAYGDDKYLRQLESEVPVYYYGVSEBDDIQARNIQRTTEGSSF 248

Qy  244 DVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
         |||:.|.| .HF:||:.|.||::NAL.VIA::Y:EKLD::::::: :.FGGVKRRF:|::
Db  249 DVYHKDDFVGHFVLPAFGHNIMNALGVIAVAYFEKLDMQKVABEMLSFKGVKRRFSEKK 308

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
          : :::VDDYAHHP:|| |:|:.|| KYP:KE::AVFQPHTF:RT:|:::|FAE:L |||
Db  309 VSDMIVDDYAHHPAEIKATIDGARQKYPDKEIIAVFQPHTFRTIALMDEFAEALDLAD 368

Qy  364 RVFLCEIFGSIRENTGALTIQDLIDKIE-GASLINEDSINVLEQFDNAVILFMGAGDIQK 422
          VFLC:IFGS RE TG.::I:DL::KI:  .:::|ED:::  ::DF:NAV::FMGAGD:QK
Db  369 EVFLCNIFGSARETQGEVRIEDLGEKIQKGGQVITEDNVSPLLDFENAVVVFMGAGDVQK 428

Qy  423 LQNAY 427
         :|:||
Db  429 FEQAY 433
```

RESULT 2
Q88WZ5
ID   Q88WZ5   PRELIMINARY;   PRT;   436 AA.
AC   Q88WZ5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN   MURC OR LP_1462.
OS   Lactobacillus plantarum.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1590;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIMB 8826 / WCFS1;
RX   MEDLINE=22480296; PubMed=12566566;
RA   Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA   Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA   Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA   Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA   De Vos W.M., Siezen R.J.;
RT   "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR   EMBL; AL935256; CAD63921.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.

KW   Ligase; Complete proteome.
SQ   SEQUENCE   436 AA; 48750 MW; 25A582DB50238CD4 CRC64;

```
Query Match             56.5%;
Best Local Similarity   55.7%;    Score 1284.5;  DB 16;  Length 436;
Matches 240; Conservative 75; Mismatches 115;  Indels 1; Pred. No. 7.6e-82;  Gaps 1;

Qy    2 THYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKED 61
          :|:|||||||||||||: :|:|:||:|||||:|| ||| |:::.|:| :||||:::|:|
Db    5 TVYYFVGIKGSGMSSLALILHDKGYVEGSDIEQYTFTQKGLAAAGIKMLPFSEDNIREG 64

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
          .||.||:|  :||||: :| :|| |||:|:||| : |:.|||:.|:|.||||||||||||
Db   65 LTVIAGNSFTDDHPEIKKAREMGLPVYRYHEFLGKLMEGFTSIGVAGTHGKTSTTGLLSH 124

Qy  122 VMNGDKKTSFLIGDGTGMGLPESDYPAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKD 181
          |::|  ::TSYLIGDG|G G P: | :.FEADEYRRHF::Y PDYAIM :DFDHPDY:KD
Db  125 VLSHIAPTSYLIGDGYGKGTPDARFVFEADEYRRHFVAYHPDYAIMNVDFDHPDYYKD 184

Qy  182 INDVFDAFQEMARHVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITDKGT 241
          :.||.|||: .::|VKKGI.AWGDD::LR::: DVP|YYYG :: DD::| NIQ ::KG:
Db  185 LADVQSAFQQFGNQVKKGIFAWGDDESLRHLDVDTPIYYYGTNRDRDFQAVNIKRTTKGS 244

Qy  242 AFDVVVDGEFYDHFLSPQYGDHTVLNALAVTAISYLEKLDVTNIKEALETFGGVKRRFNE 301
          :|:|  ::|.  .F:.  :|:|:VLN::AV:A:SY:EK:::.:|::: :.FGGVKRRF:E
Db  245 SFEVKYHDESLGKFBIPLFGEHNVLNSTAVIAVSYFEKVNLDEIRRELLDFSGVKRRFSE 304

Qy  302 TTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSK 361
          :: : ::::DDYAHHP :| || :.||:KYP:KE::AVFQPHTFSRT:|:::.|:|:SLSK
Db  305 HQVGDMVMIDDYAHHPSBIKATLDAARQKYPDKEILAVFQPHTFSRTKALMDGFAASLSK 364

Qy  362 ADRVFLCEIFGSIRENTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDI 420
          ||.VFL :IF S RE :G.::::DL:.K:  :..::.|:DS:: ::.:NAV::FMGAGDI
Db  365 ADHVFLTNIFSSAREKSGDVSSKDLAAKLPNGGEIITTDDMSALTAYHNAVAVFMGAGDI 424

Qy  421 QKLQNAYLDKL 431
         ||:::|:  ::
Db  425 QKYEKIYEDQM 435
```

RESULT 3
Q8DY77
ID   Q8DY77   PRELIMINARY;   PRT;   443 AA.
AC   Q8DY77;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase.
GN   MURC OR SAG1615.
OS   Streptococcus agalactiae (serotype V).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=216466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2603 V/R / Serotype V;
RX   MEDLINE=22222988; PubMed=12200547;
RA   Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA   Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA   Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA   DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA   Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA   Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA   Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA   Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative genomic analysis of an
RT   emerging human pathogen, serotype V Streptococcus agalactiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR   EMBL; AE014265; AAN00479.1; -.
DR   TIGR; SAG1615; -.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.

RA   Kunst F.;
RT   "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT   invasive neonatal disease.";
RL   Mol. Microbiol 45:1499-1513 (2002).
DR   EMBL; AL766852; CAD47323.1; -.
DR   SagaList; gbs1664; -.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR001706; Ribosomal_L35.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
DR   PROSITE; PS00936; RIBOSOMAL_L35; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   443 AA; 49836 MW; 0BDBDCDBC62BBD9 CRC64;

Query Match   50.7%; Score 1153.5; DB 16; Length 443;
Best Local Similarity   51.0%; Pred. No. 1.2e-72;
Matches 220; Conservative 82; Mismatches 126; Indels 3; Gaps 3;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
            ||:||||||||||:||:|||||||||||:|||:|||: |:|: ||||| |||||:||
Db   5 YHFIGIKGSGMSALALMLHQMGHNVQGSDVDKYFTQRGLEQAGVTILPFSPNNISEDLE 64

Qy   64 VIQGNAF-ASSHEEI---TVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV 122
           :|| ||| |:| ::||| :|: |||||: ||||||| ||||
Db   65 IIAGNAFRPDNNEELAYVIEKGYHFKRYHEFLGDFMRQFTSLGVAGAHGKTSTTGLLAHV 124

Qy   123 MNGDKKTSFLIGDGTGMGLPESDYFAFRACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDI 182
           :| ||:|||:|||||:|: |::|| :::| ||||||: | ::||:|||||||||: ||
Db   125 LKNITDTSFLIGDTGRGSANANYFVFEADEYERHFMPYHPEYSIITNIDFDHPDYFTGL 184

Qy   183 NDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKOSDDIYAQNIQITDKGTA 242
           || | || | : | ::| |:|::|| :::|||||||||| |::|: :|:|:||:| |
Db   185 EDVFNAFNDYAKQVQKGLFIYGBDSKLHEITSKAPIYYYGFEDSNDFIAKDITRIVNGSD 244

Qy   243 FDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETPGGVKRRFNET 302
           ||: ||| |::|::: ||:||||:|||||: :|:||:|::||:|| ||||:|||||
Db   245 FKVFYNQEIGQFHVPAYGKHNILNATAVIANLYIMGIDMALVAEHLKTFSGVKRRFTEK 304

Qy   303 TIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKA 362
           |:|:||:||||||||: ||||:|||||:|||:|| ||||||||::|||:|||:||:||
Db   305 IIDDTVIIDDFAHHPTEIIATLDAARQKYPSKEIVAIFQPHTFTRTIALLDDFAHALSQA 364

Qy   363 DRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDI 420
           | : |:|||:|| ||||||||||:||: |||:||||||:||:|||:||||||:||||
Db   365 DSVYLAQIYGSAREVDNGEVKVEDLAAKIVKHSDLVTVENVSPLLNHDNAVYVFMGAGDI 424

Qy   421 QKLQNAYLDKL 431
           |:||||||||
Db   425 QLYERSFEELL 435

RESULT 5
Q8DSP4
ID   Q8DSP4   PRELIMINARY;   PRT;   452 AA.
AC   Q8DSP4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative UDP-N-acetyl muramate-alanine ligase (EC 6.3.2.8).
GN   MURC OR SMU.1731.
OS   Streptococcus mutans.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR001706; Ribosomal_L35.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
DR   PROSITE; PS00936; RIBOSOMAL_L35; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   443 AA; 49852 MW; C1D6FC7DE6017228 CRC64;

Query Match   51.1%; Score 1163.5; DB 16; Length 443;
Best Local Similarity   51.3%; Pred. No. 2.4e-73;
Matches 223; Conservative 81; Mismatches 120; Indels 11; Gaps 4;

Qy   4 YHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMV 63
            ||:||||||||||:||:||||||||||||:|||:|||: |:|: ||||| |||||:||
Db   5 YHFIGIKGSGMSALALMLHQMGHNVQGSDVDKYFTQRGLEQAGVTILPFSPNNISEDLE 64

Qy   64 VIQGNAFASSHEE-----IVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGL 118
           :||| ||: :::   :: |:| ::||| :|: |||||: |||||||||||:
Db   65 IIAGNAFRPDNNEELAYVIEKGYQFK---RYHEFLGDFMRQFTSLGVAGAHGKTSTTGL 120

Qy   119 LSHVMNGDKKTSFLIGDGTGMGLPESDYFAFACEYRRHFLSYKPDYAIMTNIDFDHPDY 178
           |:|: :| ||:|||:|||||:|: |::|| :::| ||||||: | ::||:||||||||
Db   121 LAHVLKNITDTSFLIGDTGRGSANANYFVFEADEYERHFMPYHPEYSIITNIDFDHPDY 180

Qy   179 FKDINVDFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQITD 238
           | : |:| || | : | ::| |:|::|| :::|||||||||| |:||: :|:|:||
Db   181 FTGLEDVFNAFNDYAKQVQKGLFIYGEDPKLHEITSEAPIYYYGFEDSNDFIAKDITRTV 240

Qy   239 KGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETPGGVKRR 298
           | ||: ||| |::|::: ||:||||:|||||: :|:||:|::||:|| ||||:|||
Db   241 NGSDFKVFYNQEIGQFHVPAYGKHNILNATAVIANLYIMGIDMALVAEHLKTFSGVKRR 300

Qy   299 FNETTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAES 358
           | |: |:|:||:|||||||: ||||:|||||:|||:|| |||||||:::|||:|||:||
Db   301 FTEKIIDDTVIIDDFAHHPTEIIATLDAARQKYPSKEIVAIFQPHTFTRTIALLDEFAHA 360

Qy   359 LSKADRVFLCEIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMG 416
           ||:|| : |:|||:|| ||||||||||:||: |||:||||||:||:|||:||||||
Db   361 LSQADSVYLAQIYGSAREVDNGEVKVEDLAAKIVKHSDLVTVENVSPLLNHDNAVYVFMG 420

Qy   417 AGDIQKLQNAYLDKL 431
           ||||:||||||||
Db   421 AGDIQLYERSFEELL 435

RESULT 4
Q8E3U2
ID   Q8E3U2   PRELIMINARY;   PRT;   443 AA.
AC   Q8E3U2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   MURC OR GBS1664.
OS   Streptococcus agalactiae (serotype III).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=216495;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEM316 / Serotype III;
RX   MEDLINE=12354221; PubMed=12354221;
RA   Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA   Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,

```
OC   Streptococcus.
OX   NCBI_TaxID=1309;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UA159 / ATCC 700610 / Serotype C;
RX   MEDLINE=22295063; PubMed=12397186;
RA   Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA   Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA   Li S., Zhu H., Najar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT   "Genome sequence of streptococcus mutans UA159, a cariogenic dental
RT   pathogen.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR   EMBL; AE015001; AAN59364.1; -.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   InterPro; IPR001706; Ribosomal_L35.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
DR   PROSITE; PS00936; RIBOSOMAL_L35; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE  452 AA;  50903 MW;  4AD29B2CCA6576CE CRC64;

Query Match       49.8%;
Best Local Similarity   50.7%;   Pred. No. 3.7e-71;   DB 16;  Length 452;
Matches  218;  Conservative   84;  Mismatches  125;  Indels    3;  Gaps    3;

Qy    5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFDANNIKEDMVV 64
          |:|||||||||:|||||||| |||||||||| |||||||||||| |||  |||||||||||
Db   15 HFIGIKGSGMSALALLLHQMGVKVQGSDVDKYYFTQHGLEKAGIPILPFAESNITNDMEI 74

Qy   65 IQGNAFASSHE-EIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHVM 123
          ||||||  |:| |:|  :: ||:: |:|: ::|||||::|:|:|||||||||||||||:||
Db   75 IAGNAFRKDNNIEVAYALENGYHFKRYHEFLGEFMNQFTSLGVAGAHGKTSTTGLLAHVL 134

Qy  124 NGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHPDYFKDIN 183
          |  |:| ||||||||:||| :|||:||  :|||||: |:|:| |:|||||||||||||  |
Db  135 KNITDTSYLIGDGTGRGLANSQYFVFESDEYERHFMPYHPEYSIITNIDFDHPDYFTGVD 194

Qy  184 DVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPIYYYGFKDSDDIYAQNIQTDKGTAF 243
          |||  |    |:::|| ::  | :| |:: :::|||||||||:||  |||||| :  : |
Db  195 DVFAAFNDYAKQVKKGLFVYGERDPYLRKLTSSAPIYYYGFKDNDDFVAYDIMRSTNGSDF 254

Qy  244 DVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETT 303
          :|:  |:: : ||   :|:| |||| |: |: :|:|: |:|:||:||| ||||||| :
Db  255 KVRHGQNELGSFHVPAFGRHNVNATVIANLYIAGVEMDLVRQHLKTFSGVKRRFSEKL 314

Qy  304 IANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKAD 363
          || :::|||:|||| ||  |:  |::||| :|:|| |||||| ||  ||::|| :: :|
Db  315 INDVIIDDFAHHPTEIIATLDAARQKYPSKEIVAIFQPHTFTRTIALLDEFAQALNEAD 374

Qy  364 RVFLCRIFGSIRE-NTGALTIQDLIDKI-EGASLINEDSINVLEQFDNAVILFMGAGDIQ 421
          |:||  |:||:|| :| | : :|| |:|  ::  ::  : :::  ||:||: |||||||||
Db  375 SVYLAQIYGSAREVDHGDVKVEDLADKIVKPAKVITVDNVSPLLDHHNAVYVFMGAGDIQ 434

Qy  422 KLQNAYLDKL 431
          || :| :: :|
Db  435 LYERSFELL 444

RESULT 6
Q899J1
ID   Q899J1        PRELIMINARY;       PRT;    456 AA.
AC   Q899J1;
```

```
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN   CTC00185.
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Massachusetts / E88;
RX   MEDLINE=22457253; PubMed=12552129;
RA   Bruggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,
RA   Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA   Gottschalk G.;
RT   "The genome sequence of Clostridium tetani, the causative agent of
RT   tetanus disease.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR   EMBL; AE015936; AAO34835.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE  456 AA;  50816 MW;  31C71A197188CBB5 CRC64;

Query Match       27.8%;
Best Local Similarity   35.1%;   Pred. No. 3.8e-36;   DB 16;  Length 456;
Matches  158;  Conservative   91;  Mismatches  168;  Indels   33;  Gaps   15;

Qy    5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKEDMV 63
          :||||||||||| ||||  ||| |||||   ||| |||  : :: |   | |::|| :
Db   13 HFIGIGISMSGMAEILLKKGYKVSGSDSTKSPIID-KLINLGAEIYIGHKAENIKNVDL 71

Qy   64 VIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYT-SVAVTGAHGKTSTTGLLSHV 122
          :  ::|: :: || :  :  :||:  |||||::|:||||| ::::
Db   72 IVYTAAVSEDNPELTKALTNNIKIMNRAEFLGYLMDGHKYNIAVSGTHGKTTTTSMMSHI 131

Qy  123 -MNGDKKTSFLIGD--------GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNID 172
          | ||:|  |||           :|:|||| :|||| |:: |||:| | ||:|||:||||
Db  132 TVNANLDPTIIVGGELNIINGNVRTG---KSEYFLTACEYKESFLKFFPYIGVILNID 187

Qy  173 FDHPDYFKDINDVFDAFQEMAHNVKKG--IIAWGDDEHLRKI--EADVPIYYYGFKDSDD 228
          ||||||||:|:|::||: |:|||::||    |: :|::|  |:   ||||||||||||:||
Db  188 FDHPDYFKDINHIKNAFSKFANLIPKDGYLIACAEDRNINDIIKNIDCTIITYGL-NKGD 246

Qy  229 IYAQNIQITDKGTA-FDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKE 287
          |:|::|  :| : | |||: ::: ||||   :: |:||||||| |  :  |:|::|:|:::
Db  247 IQAKNITFDNKGCASFDVIKNSSVLFSVKLNVPGSYNVLNALASICVSFALNIDKESIIK 306

Qy  288 ALETFGGVKRRFNETTIANQV-IVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFS 346
          .|| | |::|||  | |||:|  :||  |||| |:::  ::: :|||  :  :|||||: |
Db  307 ALETFGGVKRRFELKGVRNDVTVIEDTAHHPTEIKATL-SAAKNYPSNRILCVFQPHYYT 365

Qy  347 RTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKI--EGASLIN----EDS 400
          ||::|  || :|:    |:| : |::| ::|||| ||||||:|   ::| |:|    |:|
Db  366 RTYSLFEDFTESFYNTDTLLADIYPAREKOTGIVSSDMLGNKLRKKNINCINLHSFEDI 425

Qy  401 INVL-EQFDNAVILFMGAGDIQKLQNAYL 428
          |:||  :|||||:||||||||: || ||:|
Db  426 SDYLKKETKPGDLVLIMGAGDIYKAGDLFL 455

RESULT 7
Q9RNM7
ID   Q9RNM7        PRELIMINARY;       PRT;    486 AA.
AC   Q9RNM7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
```

DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate-alanine ligase.
GN   MURC.
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZM4;
RA   Um H.W., Kang H.S.;
RT   "Sequence analysis of 41E10 fosmid clone of Zymomonas mobilis.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179611; AAD53935.1; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase.
SQ   SEQUENCE   486 AA; 52602 MW; 7DB255E70F47CCD4 CRC64;

Query Match   23.6%; Score 536.5; DB 2; Length 486;
Best Local Similarity   32.5%; Pred. No. 2.7e-29;
Matches 151; Conservative   75; Mismatches 174; Indels   65; Gaps   14;

Qy   5 HFVGIKGSGMSSLAQIMHDLGHEVQGSDI-ENVVFTEVALRNKGIKIL-PFDANNIKEDM 62
            :|||| ||| :|:::|||::||| ||::| :|||:|:||:|| :|:| |:|:||:|
Db   12 HFVFGIGIGMSGIAEVMENLGYVQGSDISEGY--TVDALRQMGIKVLIGHHAENVKDAA 69

Qy   63 VVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLS-- 120
            || : :| ||| || :||||||||||||||||:::|||||||||||||||||||||
Db   70 VVVVSAIHRGNPEVEAALENRIPVVRRAEMLAELMRLKSTVAVAGTHGKTTTTSMVAAL 129

Qy   121 -------HVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNID 172
                   |:||||:|||| ||:||| :||||||||||||||||||||:||||||||
Db   130 LDAGGIDPTVINGGIINSY----GSNARLGDSDWMVVEADESDGSFLRLDGTLAIVTNID 185

Qy   173 FDHPDYFKDINDVFDAFQEMAHNVKKGIIAWGDDEH------LRKIEADVPIYYYGFKDS 226
            |:|||:| ||::|| :|||:|| |: | |:|: |:|||||: : : | :
Db   186 PEHLDHYGSFEKVQDADAFVEFVSNVPPYGAAFLCIDHPVVQSILPRIR-DRRLITYGFSAQ 244

Qy   227 DDIYAQNIQITDKGTAFDVYV--DGEF-YDHFLSPQYGDHTVLNALAVIAISYLEKLD 281
            ||| ||||:|| ||||:||| ||:| ||||:||||||||:||||||||:|:|:|||
Db   245 ADIRAVDITPISGGNRFTAVIRGRDGDIRRIENIFLPMPGRHNIQNALSAIGVALEFSIP 304

Qy   282 VTNIKEALETFGGVKRRFN--ETTI-ANQVIVDDYAHHPREISATIETARKKYPHKEV 336
            |:|| | ||||||:||||| : |||||||||:|||||||||||||||:: |||
Db   305 DAAIRDGFARFGGVKRRFSRVGEIKICDGSVLVIDDYGHHPVEIKAVLSAARES-AQQRV 363

Qy   337 VAVFQPHTFSRTQAFLNEPAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEG---- 392
            : :|| |||| || || | : :|||| :: | : : |:||||| ::||||:
Db   364 IAVVQPHRFLRLHDLMTFQSAFNDADMVFVAPYYAAGEQPIAGVDSQALVSGLKQHGHR 423

Qy   393 --------ASLINEDSINVLEQFDNAVILFMGAGDIQK 422
                    ||:||:|:|::|:||:||:|||::||||||
Db   424 SAQAVDSPRSLAIALADIIKADDI-----------IICVGAGDIIK 458

RESULT 8
Q9F1M9
ID   Q9F1M9   PRELIMINARY;   PRT;   494 AA.
AC   Q9F1M9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MurC.
GN   MURC.
OS   Shewanella violacea.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=60217;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ishii A., Nakasone K., Sato T., Sugai M., Wachi M., Nagai K., Kato C.;
RT   "Isolation and characterization of dcw gene cluster for cell division
RT   and cell wall synthesis from a deep-sea piezophilic Shewanella
RT   violacea.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB052554; BAB19203.1; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
SQ   SEQUENCE   494 AA; 54098 MW; 96E45B6A503B48E2 CRC64;

Query Match   23.0%; Score 523.5; DB 2; Length 494;
Best Local Similarity   31.8%; Pred. No. 2.2e-28;
Matches 145; Conservative   88; Mismatches 178; Indels   45; Gaps   13;

Qy   3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKED 61
            : |||| ||| :|:::||| :|::||| |: |:||::|| ||:|: ||||||:|||
Db   34 HIYFVGIGGAGMCGIAEVLVNEGYRLSGSDIAENAVTE-RLKSLGVQIIHGHHADQVHGT 92

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
            : |||| ||| || :||||||||||||||||::|||||||||||||||||||||||:|
Db   93 DVVVVSTAIDAENPELVAAKELRIPVVQRAEMLAELMRYRHGVAVAGTHGKTTTTSLIAS 152

Qy   122 V-MNGDKKTSFLIGD-----GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
            : ::|||:|||||| ||:||| :|||||||| :|:||:|:: |||||:||| |:
Db   153 IYAQADRDPTFVIGGLLNSAGTNARLGNSRYLIAEADESDASFLHLQPMVSVVTNIEADH 212

Qy   176 PD-YFKDINDVFDAFQEMAHNVK-KGI-IAWGDDEHLRKIEADV--PIYYYGFKDSDDIY 230
            || |||||:|| |:|:|:|::| :|| | |:|| :|:|:||:| ||::| |:||:|:
Db   213 MDTYGGDFEKLKSTFIDFLHNLPFYGVAVMCIDDPVVRELLPKVGRKIVTYGFSEDADIQ 272

Qy   231 AQNIQITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALE 290
            | |||||| |||:||||||:| |||:||||||||:|||||| |::||||:|| ::|||||
Db   273 ALNFVQQGYSSHFTLRRDGVEDIAVMVNLPGEHNVLNALASIAVATEDEIEDEAIVLALA 332

Qy   291 TPFGGVKRRFNE----TTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTF 345
            |||||||:|||: |||||:|||:|||:||:||||||||||||||:: |:|:|:|
Db   333 QFPEGIGRRFQQLGTFATSKGEVMLVDDYGHHPSEVAAIIKAARLGWPDKRLVMIYQPHRY 392

Qy   346 SRTQAFLNEFAESLSKADRVFLCEIFG---------SIRENTGALTIQDLIDK 389
            ||:|||||| |:||||| :: | : |: |||:||||:|:|||:||
Db   393 SRTRDLYDDFVEVLSQVDCLILLDVYSAGEAPVPGADSRALCRSIRQ-----RGQLDP 445

Qy   390 IEGASLINEDSINVLEQF--DNAVILFMGAGDIQKL 423
            | | || :|:||:|:||: :|:||:|:|||| ||:|
Db   446 IFVAG--TBQLLSLLPDVLQDGDLLLTQAGNIGTL 479

RESULT 9
Q8E9P8
ID   Q8E9P8   PRELIMINARY;   PRT;   488 AA.
AC   Q8E9P8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE UDP-N-acetylmuramate--alanine ligase.
GN MURC OR SO4218.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Gill J., Venter J.C., Nealson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015855; AAN57190.1; -.
DR TIGR; SO4218; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR InterPro; IPR005758; MurC.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR Pfam; PF01225; Mur_ligase; 1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01082; murC; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE   488 AA; 53168 MW; DC45914D4869D353 CRC64;

Query Match   22.9%; Score 521.5; DB 16; Length 488;
Best Local Similarity   32.6%; Pred. No. 3e-28;
Matches 146; Conservative 87; Mismatches 182; Indels 33; Gaps 13;

Qy   5 HFVGIKGSGMSSLAQIMHDLGHEVGSDI-ENYVFTEVALRNKGIKI-LPFDANNIKEDM 62
            |||||||||:||:|||||||:|::| ||:|||||||||||| :|| |||||||||||
Db   24 HFVGLGGAGMGGIAEVLVNEGYVVSGSDIAQNAVTDRLCL--LGAKIHIGHGADNVQQAD 81

Qy   63 VVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV 122
           |: |||||:|||||:|:|::|||:|::: :|||||||||||||||||||||||| ::
Db   82 VVVVSTAINPQNPEIIAAKELRIPIVRRAEMLAELMRYRHGVAIAGTHGKTTTSLIASL 141

Qy   123 M-NGDKKTSFLIGD------GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDHP 176
          | ||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   142 YGQAGRDPTFVIGGLLNSAGTNARLGTSRYLIAEADESDASFLHLQPMVSVVTNIRADHM 201

Qy   177 D-YFKDINDVFDAFQEMAHNVK-KGI-IAWGDDEHLRKIEADVP--IYYYGFKDSDDIYA 231
          | |||||||||||||||||||| ||| |:||||||||||||: ::|:|||||||||||
Db   202 DTYGGDFEKLKSTFVDFLHRNLPFYGVAVVCIDDPVVREIMPRISRHIVTYGFRDDADVQA 261

Qy   232 QNIQITDKGTAFDVYVDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKRALET 291
          |:||||||||||||||||||||||||| :|||||||||||:|::::||:||:||:|
Db   262 LNFSQQGHQCRFTVRRKGKEDLDLLNLPGQHNVLNALAAIRVATEDEIDDSAIIQALAB 321

Qy   292 FGGVKRRFNE-----TTIANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTFS 346
           |||:|||| |||||||||||||||||||||||||:||:||||||:|||
Db   322 FQGIGRRFQHLGKFATPKGEVMLVDDVDYGHHPSEVAATIKAARAGWPEKRLVMAYQPHRYT 381

Qy   347 RTQAFLNEFAESLSKADRVFLCEIFGS-------IRENTGALTIQDLIDKIEGAS--L 395
          |||||||||||||:||||||||||||| |||||||||||||||||:|||:
Db   382 RTRDLIYEDFIEVLSQVDCLLLLDVYSAGEAPIPGADGRALCRSIRLRGQLDPIFIASPEQ 441

Qy   396 INEDSINVLEQFDNAVILFMGAGDIQKL 423
          |||| |||||:||||||||||| |:::|
Db   442 LAEVLPDVLQEGD--LLLTQGAGNIGAL 467

RESULT 10
Q7WFS3

ID Q7WFS3;   PRELIMINARY;   PRT;   468 AA.
AC Q7WFS3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN MURC OR BB4197.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M.T., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640449; CAE34561.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE   468 AA; 49404 MW; O45E233D594FECFD CRC64;

Query Match   22.3%; Score 506.5; DB 16; Length 468;
Best Local Similarity   31.7%; Pred. No. 3.2e-27;
Matches 145; Conservative 75; Mismatches 195; Indels 43; Gaps 14;

Qy   3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKED 61
           |:|||||||||||||||||||||||||||||:| ||||||||||||| |||||||||:||
Db   7 HIHFVGVGGSGMSGIAEVLNLGYTISGSDLNESAVTR-RLAELGMRIAIGHDRANVAGA 65

Qy   62 MVVIQGNAFASSHEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
          :::||||||:||||:|:|::||||||:::|||||||||||||||||||||||||:
Db   66 GAIVTSTAVAGDNPEVLAARAARIPVVPRAVMLAELMRLKRGIAVAGTHGKTTTSLVAS 125

Qy   122 VM-NGDKKTSFLIGD------GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
          || ||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   126 VLAAGGLDPTFVIGGRLTSAGANARLGQGEYIVVEADESDASFLNLLPVMALVTNIDADH 185

Qy   176 PD-YFKDINDVFDAFQEMAHNVK--KGIIAWGDDEHLRKIEADV--PIYYYGFKDSDDIY 230
          | |||||||||||||||||||| |||||||||||||||||||: |:|||||||||||
Db   186 MDTYGHDVARLKSAFIEFTQRLPFYGSAILCADDANVREIMPFVSRPITTYGLSPDAQVC 245

Qy   231 AQNIQITDKGTAFDVVYDGEFYDHFLSPQY------GDHTVLNALAVIAISYLEKLDVTN 284
          ||||||||||||||:|||||||||||||| ||||||||||||||:||::||||
Db   246 AQDVQA--DGTRMRFTVQRRDRDVVLPALQVELNLPGLHNVRNALAAIAVATELGVDDAA 303

Qy   285 IKEALETFGGVKRRFNE-----TTIANQ----VIVDDYAHHPREISATIETARKKYPHKE 335
          |:|||||||||:|||| |||||| ||||||||||||||:|||||:||
Db   304 IREALAAFKGVGRRFTQWGDLPVPAAHGGGTTLVDDYGHHPVEMAATLAAARGAWPQRR 363

Qy   336 VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGS-------IRENTGALTIQDL 386
          ||||||||| :| :|:||||||:||||| ||:|||||| ||||||||||||
Db   364 IVLAFQPHRYTKTRDCPEDFVRVLGSADGVLLTEVVAAGEAPLVAADGRALSRALRVAGK 423

Qy   387 IDK-IEGASLINEDSINVLEQFDNAVILFMGAGDIQK 422
          |:: ||||:|||||||||||||||||||||||||||:
Db   424 VEPVFVEDVGELPQAILDFVR--DGDVVVMGAGSISK 459

RESULT 11
Q7W4B5
ID   Q7W4B5   PRELIMINARY;   PRT;   468 AA.
AC   Q7W4B5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN   MURC OR BPP3751
OS   Bordetella parapertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=12822 / ATCC BAA-587;
RX   MEDLINE=22827954; PubMed=12910271;
RA   Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA   Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA   Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA   Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA   Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA   Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA   Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA   Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA   Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA   Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT   "Comparative analysis of the genome sequences of Bordetella pertussis,
RT   Bordetella parapertussis and Bordetella bronchiseptica.";
RL   Nat. Genet. 35:32-40(2003).
DR   EMBL; BX640434; CAE43034.1; -.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   468 AA; 49404 MW; 045E233D594FECFD CRC64;

Query Match   22.3%; Score 506.5; DB 16; Length 468;
Best Local Similarity 31.7%; Pred. No. 3.2e-27;
Matches 145; Conservative   75; Mismatches 195; Indels 43; Gaps 14;

Qy   3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKED 61
          |||||:|||||:|||| :||||||||:|:|: :||: |||:||||:||| :||||||||
Db   7 HIHFVGVGGSGMSGIAEVLLNLGYTISGSDLNESAVTR-RLAELGMRIAIGHDRANVAGA 65

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
          : |::|||||||::: | |: || ::::|:|| :: ||||||||||||||||||||||||
Db   66 GAIVTSTAVAGDNPEVLAARAARIPVVPRAVMLAELMRLKRGIAVAGTHGKTTTTSLVAS 125

Qy   122 VM-NGDKKTSFLIGD------GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
          |: ||::| |||| | ||:: |::| ::|| :::| :|||||:||:||
Db   126 VLAAGGLDPTFVIGGRLTSAGANARLGQGEYIVVEADESDASFLNLLPVMAIVTNIDADH 185

Qy   176 PD-YFKDINDVFDAFQEMAHNVK--KGIIAWGDDEHLRKIEADV--PIYYYGFKDSDDIY 230
          | |: | | |||:::||| |: ||::| |::|| ||::||| :|| ||:|||||||
Db   186 MDTYGHDVARLKSAFIEFTQRLPFYGSAILCADDANVREIMPFVSRPITTYGLSPDAQVC 245

Qy   231 AQNIQITDKGTAFDVVYDGEFYDHFLSPQY------GDHTVLNALAVIAISYLEKLDVTN 284
          |::| ||:|| |:||::| |::|| ||| ||||::|| :::|: |: |::| ||||
Db   246 AQDVQA--DGTRMRFTVQRRDRDVVLPALQVELNLPGLHNVRNLALAAIAVATELGVDDAA 303

Qy   285 IKEALETFPGVKRRFNE-----TTIANQ---VIVDDYAHHPREISATIETARKKYPHKE 335
          |||| ||||:|:||:|:   ::|:   :||::|:||:|:||::|:|||:||:||
Db   304 IREALAAFKGVGRRFTQWGDLPVPAAHGGGIFTLVDDYGHHPVEMAATLAAARGAWPQRR 363

Qy   336 VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGS-------IRENTGALTIQDL 386
          :|| ||||:|:|||| |:| ::| |:||||:|||||| |||||||:|||||
Db   364 IVLAFQPHRYTRTRDCFEDFVRVLGSADGVLLTEVYAAGEAPLVAADGRALSRALRVAGK 423

Qy   387 IDK--IEGASLINEDSINVLEQFDNAVILFMGAGDIQK 422
          ||| |||:||||||||||:|:|||||:|||||||::|
Db   424 VEPVFVEDVGELPQAILDFVR--DGDVVVMGAGSISK 459

RESULT 12
Q7VUQ4
ID   Q7VUQ4   PRELIMINARY;   PRT;   468 AA.
AC   Q7VUQ4;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8).
GN   MURC OR BP3022.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=520;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX   MEDLINE=22827954; PubMed=12910271;
RA   Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA   Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA   Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA   Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA   Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA   Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA   Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA   Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA   Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA   Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT   "Comparative analysis of the genome sequences of Bordetella pertussis,
RT   Bordetella parapertussis and Bordetella bronchiseptica.";
RL   Nat. Genet. 35:32-40(2003).
DR   EMBL; BX640420; CAE43293.1; -.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   468 AA; 49416 MW; 1E77233D4366E724 CRC64;

Query Match   22.3%; Score 506.5; DB 16; Length 468;
Best Local Similarity 31.7%; Pred. No. 3.2e-27;
Matches 145; Conservative   75; Mismatches 195; Indels 43; Gaps 14;

Qy   3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKI-LPFDANNIKED 61
          |||||:|||||:|||| :||||||||:|:|: :||: |||:||||:||| :||||||||
Db   7 HIHFVGVGGSGMSGIAEVLLNLGYTISGSDLNESAVTR-RLAELGMRIAIGHDRANVAGA 65

Qy   62 MVVIQGNAFASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
          : |::|||||||::: | |: || ::::|:|| :: ||||||||||||||||||||||||
Db   66 GAIVTSTAVAGDNPEVLAARAARIPVVPRAVMLAELMRLKRGIAVAGTHGKTTTTSLVAS 125

Qy   122 VM-NGDKKTSFLIGD------GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
          |: ||::| |||| | ||:: |::| ::|| :::| :|||||:||:||
Db   126 VLAAGGLDPTFVIGGRLTSAGANARLGQGEYIVVEADESDASFLNLLPVMAIVTNIDADH 185

Qy   176 PD-YFKDINDVFDAFQEMAHNVK--KGIIAWGDDEHLRKIEADV--PIYYYGFKDSDDIY 230
          | |: | | |||:::||| |: ||::| |::|| ||::||| :|| ||:|||||||
Db   186 MDTYGHDVARLKSAFIEFTQRLPFYGSAILCADDANVREIMPFVSRPITTYGLSPDAQVC 245

Qy   231 AQNIQITDKGTAFDVVYDGEFYDHFLSPQY------GDHTVLNALAVIAISYLEKLDVTN 284
          |::| ||:|| |:||::| |::|| ||| ||||::|| :::|: |: |::| ||||
Db   246 AQDVQA--DGTRMRFTVQRRDRDVVLPALQVELNLPGLHNVRNLALAAIAVATELGVDDAA 303

Qy   285 IKEALETFPGVKRRFNE-----TTIANQ---VIVDDYAHHPREISATIETARKKYPHKE 335
          |||| ||||:|:||:|:   ::|:   :||::|:||:|:||::|:|||:||:||
Db   304 IREALAAFKGVGRRFTQWGDLPVPAAHGGGIFTLVDDYGHHPVEMAATLAAARGAWPQRR 363

Qy   336 VVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGS-------IRENTGALTIQDL 386
          :|| ||||:|:|||| |:| ::| |:||||:|||||| |||||||:|||||
Db   364 IVLAFQPHRYTRTRDCFEDFVRVLGSADGVLLTEVYAAGEAPLVAADGRALSRALRVAGK 423

Qy   387 IDK--IEGASLINEDSINVLEQFDNAVILFMGAGDIQK 422
          ||| |||:||||||||||:|:|||||:|||||||::|
Db   424 VEPVFVEDVGELPQAILDFVR--DGDVVVMGAGSISK 459

RESULT 13
Q87WY6
ID   Q87WY6   PRELIMINARY;   PRT;   486 AA.
AC   Q87WY6;

Mon Jun 7 10:15:24 2004

us-10-712-713-2.rspt

```
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase.
GN   MURC OR PSPTO4407.
OS   Pseudomonas syringae (pv. tomato).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=323;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DC3000;
RA   Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA   Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA   Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA   Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA   White O., Fraser C., Collmer A.;
RT   "Complete sequence of Pseudomonas syringae.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016871; AAO57856.1; -.
DR   TIGR; PSPTO4407; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   486 AA;  52671 MW;  9020894E5F27C757 CRC64;
```

Query Match   22.1%;
Best Local Similarity   33.1%;   Pred. No. 7e-27;
Matches 154; Conservative   80; Mismatches 193; Indels   38; Gaps   16;

```
Qy      5 HFVGIKGSGMSSLLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKILPFD--ANNIKEDM 62
            ||||||| :| | ||||  | :||.| :|||.|| |:|| :|:|||||||:|    :||:|
Db     20 HFVGIGGVGMCGIAEVLLNLGYEVSGSDLKGSAVTE-RLESFGAQIFVGHRAENTVGADV 78

Qy     63 VVIQGNAFASSHERIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSHV 122
            ::|.|::| :.| || .|:|:::|:||| :|| : | | | :|||.||||.|:||:
Db     79 LVV--SSAVNTSNPEVATALERRIPVVPRAEMLAELMRYRHGIAVAGTHGKTTTSLLASV 137

Qy    123 M-NGDKKTSFLIGD----GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH- 175
            : :|:  |||.| |     ||||||||||  |:|| | |||||..|||: |||:||:
Db    138 FAAGLDPTFVIGGRINAAGTNAQLGTSRYLIAEADSDSDIHLQPIVAVVNIDADHM 197

Qy    176 PDYFKDINDVFDAFQEMAHNVK-KGI-IAWGDEHLRKIEADV--PIYYYGFKDSDDIYA 231
            |||||||||| |  |:| ||||  || |||||||||::|| |   |:.:||:||||: |
Db    198 ATYGGDFNKLKKTFVEFLHNLPFYGLAVMCIDDPVVRDILPLVKRPTMTYGFSEDSDVRA 257

Qy    232 QNIQITDKGTAFDVY-VDGEFYDHFLSPQYGDHTVLNALAVIAISYLEKLDVTNIKEALE 290
            | | | | | |||:|  |:| |:|| :|||:.|| : ||.:|: |..::
Db    258 INVRDQGMLTFFTVLRRDREPLDVSVN-MPGNHNVLNSLATTIAIATDEGVSDEAIVQGLS 316

Qy    291 TFGGVKRRF--NETTI--ANQVIVDDYAHHPREISATIETARKKYPHKEVVAVFQPHTF 345
            | |||.||:   |:||    |||||||||||||||||:| ||:||| ||:|||:||||
Db    317 GFQGVGRRFQVVGELPVEGGHVMLVDDYGHHPREVAAVISAVRGGWPDRRLVMVYQPHRF 376

Qy    346 SRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIBGASLINEDSINVLE 405
            ||:|||||||||||:||||||| ||:||||||||||||||||:|||.||:||||:|||::
Db    377 SRTRDLYDDFVQVLAEANVLLMEVYPAGEEPVPGADSRNLCHSIRQRGOL--DPIYIER 434

Qy    406 QFDNA--------VILFMGAGDIQKLQNAYLDK--LGMKNA 436
                       :|:|||||||||||||| :||.|:   |||:|
Db    435 GVELAPLVKPLLRAGDILLCQGAGDIAGLAPRILLNSPLFVGAKVA 479
```

Query Match   22.0%;   Score 500; DB 1; 9e-27; Length 465;
Best Local Similarity   29.5%;   Pred. No. 9e-27;
Matches 143; Conservative   86; Mismatches 177; Indels   78; Gaps   16;

```
Qy      1 MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALRNKGIKIL-PFDANNIK 59
            : ||:||||||||||||||||||||||||:  | | :||.| |||||||:|  |::|||:
Db      8 INHVHCLGIGCGIGVSALAEILLKKGCRVTGSDVSPNQVTE-RLQRLGAEIIFNHDGTAIT 66

Qy     60 EDMVVIQGNAFASSHEBIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLL 119
            .|  |:|.|:| :.| ||||:|:::|:||| :|| : | | | :|||.||||.|:||:
Db     67 QADCAVYSSAIGATNPELMAAKQAKIPLLKRGEMLANLMKEYQSIAVAGAHGKTTTSGML 126

Qy    120 SHV-MNGDKKTSFLIG----------DGTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDF 173
            |:|  |:|  |||:||            ||||||||||  |:|| | |||||..|||: |||:|
Db    127 SHAFVEANLDPTFMVGGVLNNSQTPARVGNGHYFIAEADESDASFLMHPDIAVVTNIDA 186

Qy    174 DH-PDYFKDINDVFDAF-QEMAHNVKKGIIAWG-DDEHLRKIBADVP--IYYYGFKDSDD 228
            || |||||||||| |  |:| ||||  || ||  |||||::|| |   |:.:||:|||
Db    187 DHLSTYDGDFNRLKQTYIQFLEQTAQDGVVVLCLDDPILREIAPLLSRVITYGF--SSD 244

Qy    229 IYAQNIQITDKGTAFDVYVDGEFYDHFLSPQY----------GDHTVLNALAVIAISYL 277
            |:|| | | | | |||:|  |:| |:|| :|||        |||: : ||.:|: |..:
Db    245 AQYRVVDYCQQG-------IQSLFQIH--SPQRKAPLTVKLSMPGQHNALNATAVTAIADV 296

Qy    278 EKLDVTNIKEALETFGGVKRRFNETTI-----ANQVIVDDYAHHPREISATIETARK 329
            |||||:|||||||:|:||:||||||::     |||||||||||||||:|||||||||
Db    297 VQMNEPALLKSLADFPGVDRRF--TIRGEMILPKGNALIIEDYGHHPNEIKATLAAARA 353

Qy    330 KYPHKEVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIF--------LGMKNA 371
            ||| ||:|||:||||:||:|||||||||||:|||||||:|:|         :|:
Db    354 AWPERMVLVFQPHRYSRTRDLMTEFVSVLAETDWLLVLEVYSAGEMPIPGADGMALIKM 413

Qy    372 --GSIRENTGALTIQDLIDKIEGASLINEDSINVLEQFQFDNAVILFMGAGDIQKLQNAYL 428
                ||:||||||||||||:|:.:  | |||:|:|:|||:|||||||:|:||||| :
Db    414 MSNGMAQKTTFVPLLQNLPETLQKLSQPND-------TIILQGAGNIGSIVTALV 461
```

```
RESULT 14
ID   Q83F17      PRELIMINARY;      PRT;   465 AA.
AC   Q83F17;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   UDP-N-acetylmuramate--alanine ligase.
GN   MURC OR CBU0136.
OS   Coxiella burnetii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Coxiellaceae; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nine Mile phase I / RSA 493;
RX   MEDLINE=22608657; PubMed=12704232;
RA   Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA   Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA   DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA   Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA   Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT   "Complete genome sequence of the Q-fever pathogen, Coxiella
RT   burnetii.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR   EMBL; AE016960; AAO89700.1; -.
DR   TIGR; CBU0136; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   465 AA;  50640 MW;  5B2953C14A59991B CRC64;
```

Qy   429 DKLG 432
          | -
Db   462 QTHG 465

Db   348 TMAAARNAFPGRRLVLAFQPHRYSRTRDLFEDFVRVLSGADVLLLTEVYPAGEPIIAAD 407
          |: || |:|:| ||||| :|||| :|:: | || |||:|: ::| |
Qy   383 IQDLIDKIEGASLINEDSINVLEQF-------DNAVILFMGAGDIQK 422
          |: || ||||||| ||||| ||||| ||||||| |||
Db   408 SKSLARAIRVQGKIEPIYIEQIDELKATIHTIAQDGDVILIMGAGSIGK 456

Search completed: June 3, 2004, 14:40:38
Job time : 48 secs

RESULT 15
Q82VS2
ID   Q82VS2   PRELIMINARY;   PRT;   473 AA.
AC   Q82VS2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MurC; UDP-N-acetylmuramate--alanine ligase protein (EC 6.3.2.8).
GN   MURC OR NE0992.
OS   Nitrosomonas europaea.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC   Nitrosomonadaceae; Nitrosomonas.
OX   NCBI_TaxID=915;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19718 / IFO 14298;
RX   MEDLINE=22586410; PubMed=12700255;
RA   Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA   Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA   Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT   "Complete genome sequence of the ammonia-oxidizing bacterium and
RT   obligate chemolithoautotroph Nitrosomonas europaea.";
RL   J. Bacteriol. 185:2759-2773 (2003).
DR   EMBL; BX321859; CAD84903.1; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0008763; F:UDP-N-acetylmuramate-L-alanine ligase activity; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR   InterPro; IPR005758; MurC.
DR   InterPro; IPR000713; Mur_ligase.
DR   InterPro; IPR004101; Mur_ligase_C.
DR   Pfam; PF01225; Mur_ligase; 1.
DR   Pfam; PF02875; Mur_ligase_C; 1.
DR   TIGRFAMs; TIGR01082; murC; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   473 AA;   51271 MW;   BE5A7F7C42471FD7 CRC64;

Query Match   22.0%; Score 500; DB 16; Length 473;
Best Local Similarity   32.2%; Pred. No. 9.3e-27;
Matches 151; Conservative 74; Mismatches 176; Indels 68; Gaps 18;

Qy   3 HYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENVVFTEVALRNKGIKI-LPFDANNIKED 61
          | ||||||||||||| |||:|: ||| ||||: |||| | :| | : ||: || :
Db   7 HIHFVGIGGSGMGGIAEVLINLGFQISGSDMHSNSTTR-RLQCLGAVIHHTHAAENIQSA 65

Qy   62 MVVIQGNAPASSHEEIVRAHQLKLDVVSYNDFLGQIIDQYTSVAVTGAHGKTSTTGLLSH 121
          |: | :|: |: || |||: || ||| :||: ||: :|:|:|| |||:|::|:
Db   66 DAVVISTAIHSDNPEVIAARERRIPVPRAMMLAELLRLRRGIAIAGTHGKTTTTSLVAS 125

Qy   122 VM-NGDKKTSFLIGD-----GTGMGLPESDYFAFEACEYRRHFLSYKPDYAIMTNIDFDH 175
          : : :|:||| |:|| :| : : ||::|| ::|||: | | ::|||||
Db   126 ILAEAGQDPTFVIGGKLKTVDSHARLGKGEFIVEADESDASFLYLQPVLTVVTNIDADH 185

Qy   176 -PDYFKDINDVFDAFQE-MAHNVKKGI-IAWGDDEHLRKIEADV--PIYYYGFKDSD-DI 229
           |||| ||| ||:| : :| ||| : || : | ::|: || ||| |:| |:
Db   186 MSTYEHDFNRLKQTFVEFIEHLPFYGMAVLCVDDPHVREIISMITRPVTTYGIASEDAQI 245

Qy   230 YAQNIQITDKGTAFDVYVDGEFYD---HPL-------SPQY--------GDHTVLNALA 270
          :| | |: : | | | |: | ||| | || | ||||||||
Db   246 CATNIR--------HDRCRMHFLAHIGVNGSPRTLEVTLNLPGKHNVLNALA 289

Qy   271 VIAISYLEKLDVTN--IKEALETFGGVKRRF---NETTIANQ--VIVDDYAHHPREISA 322
          :: || || | | : |:||:| |:|| :|| | | |: |||:| |: |
Db   290 AIAVG--NELGVPDEAIVKALATFGGVDRRFQQYGEIPLPDQGSFALIDDYGHHPAEIAA 347

Qy   323 TIETARKKYPHKEVVAVFQPHTFSRTQAPLNEFAESLSKADRVFLCEIFGSIRENTGALT 382

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June   5, 2004, 22:02:04 ; Search time 5350 Seconds
                                             (without alignments)
                                  10677.779 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacactcatttgt.........gaaaatgcgtttaagctt 1318
Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_ba:*
                2: gb_htg:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
               10: gb_ro:*
               11: gb_sts:*
               12: gb_sy:*
               13: gb_un:*
               14: gb_vi:*
               15: em_ba:*
               16: em_fun:*
               17: em_hum:*
               18: em_in:*
               19: em_mu:*
               20: em_om:*
               21: em_or:*
               22: em_ov:*
               23: em_pat:*
               24: em_ph:*
               25: em_pl:*
               26: em_ro:*
               27: em_sts:*
               28: em_un:*
               29: em_vi:*
               30: em_htg_hum:*
               31: em_htg_inv:*
               32: em_htg_other:*
               33: em_htg_mus:*
               34: em_htg_pln:*
               35: em_htg_rod:*
               36: em_htg_mam:*
               37: em_htg_vrt:*
               38: em_sy:*
               39: em_htgo_hum:*
               40: em_htgo_mus:*
               41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1314	99.7	1335	6	AX742069	AX742069 Sequence
2	1314	99.7	2424	6	AR354274	AR354274 Sequence
3	1312.4	99.6	1335	6	AX742071	AX742071 Sequence
4	1311	99.5	1332	6	AX617936	AX617936 Sequence
5	1310.8	99.5	301550	1	AP003134	AP003134 Staphyloc
6	1310.8	99.5	342600	1	AP003363	AP003363 Staphyloc
7	1309.2	99.3	272850	1	AP004828	AP004828 Staphyloc
8	1302.8	98.8	1314	6	AF034076	AF034076 Staphyloc
9	1302.8	98.8	1314	6	AX191741	AX191741 Sequence
10	1293.2	98.1	1351	6	E35613	E35613 MurC. 6/200
11	1293.2	98.1	1351	6	BD178030	BD178030 MurC. 4/2
12	918.8	69.7	300029	1	AE016748	AE016748 Staphyloc
13	637.8	48.4	660	6	E35614	E35614 MurC. 6/200
14	637.8	48.4	660	6	BD178031	BD178031 MurC. 4/2
15	597.6	45.3	291804	1	AE017039	AE017039 Bacillus
16	596	45.2	300975	1	AE017013	AE017013 Bacillus
17	570.4	43.3	619	6	AR194630	AR194630 Sequence
18	553.4	42.0	302241	1	AE016953	AE016953 Enterococ
19	545.2	41.4	299950	1	AE004600	AE004600 Oceanobac
20	523	39.7	299950	1	AL596169	AL596169 Listeria
21	523	39.7	349980	6	AX417045	AX417045 Sequence
22	517.4	39.3	349980	6	AX413017	AX413017 Sequence
23	516.4	39.2	240050	1	AL591979	AL591979 Listeria
24	516.4	39.2	349980	6	AX641669	AX641669 Sequence
25	500.8	38.0	199922	1	BSUB0016	Z99119 Bacillus su
26	500.8	38.0	220060	1	AP008220	AP008220 Bacillus
27	495.6	37.6	303250	1	AP001518	AP001518 Bacillus
28	456	34.6	10417	1	AE006433	AE006433 Lactococc
29	450.8	34.2	1332	6	AX607160	AX607160 Sequence
30	450.8	34.2	20601	1	AE014265	AE014265 Streptoco
31	450.8	34.2	44145	1	AX602195	AX602195 Sequence
32	450.8	34.2	174050	1	SAG766852	AL766852 Streptoco
33	447.8	34.0	10470	1	AE006499	AE006499 Streptoco
34	446.2	33.9	52276	1	AE014141	AE014141 Streptoco
35	446.2	33.9	323825	1	AP005146	AP005146 Streptoco
36	443	33.6	11900	1	AE009982	AE009982 Streptoco
37	434.8	33.0	302050	1	AL935256	AL935256 Lactobaci
38	426.2	32.3	12221	1	AE015001	AE015001 Streptoco
39	412.4	31.3	2320	1	BACDNAM	L31845 Bacillus su
40	402.4	30.5	10461	1	AE008507	AE008507 Streptoco
41	400.8	30.4	1332	6	AX569902	AX569902 Sequence
42	400.8	30.4	10344	1	AE007447	AE007447 Streptoco
43	400.8	30.4	11864	6	AR218829	AR218829 Sequence
44	400.8	30.4	11864	6	BD003741	BD003741 Polymucle
45	400.8	30.4	297172	2	SPNEU1905	AL449927 Streptoco

ALIGNMENTS

RESULT 1
AX742069
LOCUS   AX742069   Sequence 27 from Patent WO03025007.   1335 bp   DNA   linear   PAT 10-MAY-2003
DEFINITION   Sequence 27 from Patent WO03025007.
ACCESSION   AX742069
VERSION   AX742069.1   GI:30524566
KEYWORDS   .
SOURCE   Staphylococcus aureus
  ORGANISM   Staphylococcus aureus
              Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Edwards,A., Dharamsi,A., Vedadi,M., Alam,M.Z., Awrey,D.,
              Beattie,B., Domagala,M., Houston,S., Kanagarajah,D., Nethery,K.,
              Ng,I., Mansoury,K., Mcdonald,M.L., Pinder,B., Viola,C. and
              Wrezel,O.

TITLE   Novel purified polypeptides involved in membrane biosynthesis
JOURNAL   Patent: WO 03025007-A 27 27-MAR-2003;
         Affinium Pharmaceuticals Inc. (CA)
FEATURES   Location/Qualifiers
  source   1..1335
             /organism="Staphylococcus aureus"
             /mol_type="genomic DNA"
             /db_xref="taxon:1280"
ORIGIN

Query Match   99.7%; Score 1314; DB 6; Length 1335;
Best Local Similarity 100.0%; Pred. No. 5.5e-190;
Matches 1314; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 81

Qy   61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 141

Qy   121 GTTGCTCCTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GTTGCTCCTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 201

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAGTAGTACGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAGTAGTACGTGCA 261

Qy   241 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 321

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 381

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTCACAGGTATGGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy   421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATAATAGACGTCACTTTTTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATAATAGACGTCACTTTTTAAGT 501

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 TATAAACCTGATTACGCAATTATGACAAATATTGATCATCCTGATTATTTTAAA 561

Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGTATT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGTATT 621

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 681

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGGATAAAGGT 741

Qy   721 ACTGCTTTGACTTTGTGATGTGGATGGGAGTTTTATGATCATCCGTCTCCACAATAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGCTTTGACTTTGTGATGTGGATGGGAGTTTTATGATCACTTCCGTCTCCACAATAT 801

Qy   781 GGTGACCATACAGTTTAAATGTTTAAATGCGATTAGCTGGATTAGTTATTTTAGAGAAGCTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GGTGACCATACAGTTTTAAATGTTTAAATGCGATTAGTTGGATTAGTTATTTTAGAGAAGCTA 861

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTGGTGGTGTTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTGGTGGTGTTAAACGTCGTTTCAAT 921

Qy   901 GAAACTACACATTGCAAATCAAGTCAAGTCAAGTTATTGTTGAGATGATTATTGCCATCAAGAGAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACACATTGCAAATCAAGTCAAGTTATTGTTGAGATGATTATTGCCATCAAGAGAAATT 981

Qy   961 AGTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCTAGAACACAGGCATTTTTAAATGAATTTGCAGAAAGTTTAAGT 1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAATTTTGATCAATTAGAGAAAAATACTGGCGCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGAATTTTGATCAATTAGAGAAAAATACTGGCGCA 1161

Qy   1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCT 1221

Qy   1201 ATTAATGATTAGAACAATTTGATAATGTCGTGTTATTTTATTTATTGGTGCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGATTAGAACAATTTGATAATGTCGTGTTATTTTATTTATTGGTGCAGGTGATATT 1281

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAATGCGTTTTAA 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAATGCGTTTTAA 1335

RESULT 2
AR354274
LOCUS   AR354274   2424 bp   DNA   linear   PAT 17-AUG-2003
DEFINITION Sequence 392 from patent US 6593114.
ACCESSION   AR354274
VERSION   AR354274.1 GI:3760358
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2424)
  AUTHORS   Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE   Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6593114-A 392 15-JUL-2003;
FEATURES   Location/Qualifiers
  source   1..2424
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN

Query Match   99.7%; Score 1314; DB 6; Length 2424;
Best Local Similarity 100.0%; Pred. No. 4.6e-190;
Matches 1314; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   627 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 686

Qy   61 ATGCATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   687 ATGCATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 746

Qy   121 GTTGCTCCTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   747 GTTGCTCCTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 806

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAGTAGTACGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   807 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAGTAGTACGTGCA 866

Qy   241 CATCAATTGAAATTAGATGTGTAAGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
LOCUS        AX742071                1335 bp    DNA     linear   PAT 10-MAY-2003
DEFINITION   Sequence 29 from Patent WO03025007.
ACCESSION    AX742071
VERSION      AX742071.1  GI:30524567
KEYWORDS     .
SOURCE       Staphylococcus aureus
ORGANISM     Staphylococcus aureus
             Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE    1
AUTHORS      Edwards,A., Dharamsi,A., Vedadi,M., Alam,M.Z., Awrey,D.,
             Beattie,B., Domagala,M., Houston,S., Kanagarajah,D., Nethery,K.,
             Ng,I., Mansoury,K., Mcdonald,M.L., Pinder,B., Viola,C. and
             Wrezel,O.
TITLE        Novel purified polypeptides involved in membrane biosynthesis
JOURNAL      Patent: WO 03025007-A 29 27-MAR-2003
             Affinium Pharmaceuticals, Inc. (CA)
FEATURES     Location/Qualifiers
   source    1..1335
             /organism="Staphylococcus aureus"
             /mol_type="genomic DNA"
             /db_xref="taxon:1280"
ORIGIN

Query Match          99.6%;  Score 1312.4;  DB 6;   Length 1335;
Best Local Similarity 99.9%;  Pred. No. 9.6e-190;
Matches 1313;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      1 ATGACACACTATCATTTGTCGGAATTAAAGTTCGGCATGAGTTCATTAGCACAAATC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATTTGTCGGAATTAAAGTTCGGCATGAGTTCATTAGCACAAATC    81

Qy     61 ATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA   141

Qy    121 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTTGATGCTAATAACATAAAAGAA   201

Qy    181 GATATGGTAGTTATACAAGGTAATGGTCCATGAAGGTAATGGTGCAGTAGTACGTGCA   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GATATGGTAGTTATACAAGGTAATGGTCCATGAAGGTAATGGTGCAGTAGTACGTGCA   261

Qy    241 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA   321

Qy    301 TATACTTCAGTAGCTGTGCACATGGTCGTGCACATGGTAAAACTTCTACACAGGTTATTATCA   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTGCACATGGTCGTGCACATGGTAAAACTTCTACACAGGTTATTATCA   381

Qy    361 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGTATGGGA   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGTATGGGA   441

Qy    421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAGT   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAGT   501

Qy    481 TATAAACCTGATTACGCAATATGACAAATATTGATTCGATCATCCTGATATTTTAAA   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTACGCAATATGACAAATATTGATTCGATCATCCTGATATTTTAAA   561

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGCACATAAATGTTAAAAAAAGGTATT   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGCACATAAATGTTAAAAAAAGGTATT   621

Qy    601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTACGGAGACAGATGTTCCAATTATTAT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTACGGAGACAGATGTTCCAATTATTAT   681

Qy    661 TATGGATTAAAGATTCGGATGACATTCAAAAATATTCAAAAATATTCAAATTACCGATAAAGGT   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGATTAAAGATTCGGATGACATTCAAAAATATTCAAAAATATTCAAATTACCGATAAAGGT  741
```

RESULT 3
AX742071

```
 61 ATGCATGATTTAGGACATGATGAAGTTCAAGGATCCGGATATTGAGAACTACGTATTTACAGAA  120
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  82 ATGCATGATTTAGGACATGATGAAGTTCAAGGATCCGGATATTGAGAACTACGTATTTACAGAA  141

121 GTTGCTCTTAGAAATAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  180
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 GTTGCTCTTAGAAATAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  201

181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA  240
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA  261

241 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTAATTGATCAA  300
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 262 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTAATTGATCAA  321

301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA  360
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA  381

361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA  420
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA  441

421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGGCATGTGAATATAGACGTCACTTTTTAAGT  480
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGGCATGTGAATATAGACGTCACTTTTTAAGT  501

481 TATAAACCTGATTACGCAATTATGACAAATTTTGATCATCCTGATTATTTTTAAA  540
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 502 TATAAACCTGATTACGCAATTATGACAAATTTTGATCATCCTGATTATTTTTAAA  561

541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  600
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  621

601 ATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 622 ATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  681

661 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGATAAAGGT  720
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 682 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGATAAAGGT  741

721 ACTGCTTTTGATGTGTATGTGAGTTTATGATCACTTCCTGTCTCCGTCTCCACAATAT  780
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 742 ACTGCTTTTGATGTGTATGTGAGTTTATGATCACTTCCTGTCTCCGTCTCCACAATAT  801

781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA  840
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA  861

841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  900
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  921

901 GAAACTACAATTGCAAATCAAGTTATTGTGATGATCAGTTATGATGATTCCAAGAGAAATT  960
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 922 GAAACTACAATTGCAAATCAAGTTATTGTGATGATCAGTTATGATGATTCCAAGAGAAATT  981

961 AGTGCTACAATTGAAACAGCACGAAAGAAAATCCACATAAAGAAGTTGTTGCAGTATTT  1020
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 982 AGTGCTACAATTGAAACAGCACGAAAGAAAATCCACATAAAGAAGTTGTTGCAGTATTT  1041

1021 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  1080
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1042 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  1101

1081 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA  1140
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1102 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA  1161

1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200
Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
682 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGGATAAAGGT  741
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  780

742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  801
Qy 781 GGTGACCATACAGTTTAAATGCGATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA  840
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
802 GGTGACCATACAGTTTAAATGCGATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA  861

841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  900
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  921

901 GAAACTACAATTGCAAATCAAGTTATTGTGATGATCAGTTATGATGATTCCAAGAGAAATT  960
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
922 GAAACTACAATTGCAAATCAAGTTATTGTGATGATCAGTTATGATGATTCCAAGAGAAATT  981

961 AGTGCTACAATTGAAACAGCACGAAAGAAAATCCACATAAAGAAGTTGTTGCAGTATTT  1020
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
982 AGTGCTACAATTGAAACAGCACGAAAGAAAATCCACATAAAGAAGTTGTTGCAGTATTT  1041

1021 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  1080
Db      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
1042 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  1101

1081 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA  1140
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1102 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA  1161

1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1221

1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATATT  1260
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1222 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATATT  1281

1261 CAAAAATTACAAAAATGCATATTTAGATAATAAATTAGGCATGAGAAAAAATGCGTTTTAA  1314
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
1282 CAAAAATTACAAAAATGCATATTTAGATAATAAATTAGGCATGAGAAAAAATGCGTTTTAA  1335
```

```
RESULT 4
AX617936
LOCUS       AX617936                1332 bp    DNA       linear   PAT 20-FEB-2003
DEFINITION  Sequence 899 from Patent WO02094868.
ACCESSION   AX617936
VERSION     AX617936.1  GI:28448287
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Masignani,V.C., Mora,M.C. and Scarselli,M.C.
  TITLE     Staphylococcus aureus proteins and nucleic acids
  JOURNAL   Patent: WO 02094868-A 899 28-NOV-2002;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1..1332
                     /organism="Staphylococcus aureus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1280"
ORIGIN

Query Match     99.5%;  Score 1311;  DB 6;  Length 1332;
Best Local Similarity   100.0%;  Pred. No. 1.6e-189;
Matches 1311;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGACACCACTATCATTTTGTCGCGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACCACTATCATTTTGTCGCGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC  81
```

```
Db      1162  TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1201  ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATTGGGTGCAGTGATATT 1260
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      1222  ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATTGGGTGCAGTGATATT 1281

Qy      1261  CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAAATGCGTTT 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1282  CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAAATGCGTTT 1332
```

RESULT 5
AP003134/c
LOCUS   AP003134   301550 bp   DNA   linear   BCT 24-APR-2003
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 6/10.
ACCESSION   AP003134 BA000018
VERSION   AP003134.2 GI:14349226
KEYWORDS   .
SOURCE   Staphylococcus aureus subsp. aureus N315
ORGANISM   Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
TITLE   Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
MEDLINE   21311952
PUBMED   11418146
REFERENCE   2 (bases 1 to 301550)
AUTHORS   Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
TITLE   Direct Submission
JOURNAL   Submitted (30-JAN-2001) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation; Biotechnology
            Center, 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:biolo@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1935, Fax:81-3-3481-8424)
COMMENT   On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES   Location/Qualifiers
     source   1..301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /mol_type="genomic DNA"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene   complement(249..947)
                     /gene="SA1293"
     CDS   complement(249..947)
                     /gene="SA1293"
                     /note="ORFID:SA1293"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42553.1"
                     /db_xref="GI:13701259"
                     /translation="MSSIFTMIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKCKVVSLSPANVDRPSVAGTAIAAHEVG
                     HAIQDHQGYWPLRFRAALVPVANLGSSLSYMLIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILIARSSD"
     gene   complement(1001..1588)
                     /gene="SA1294"
     CDS   complement(1001..1588)
                     /gene="SA1294"
                     /note="ORFID:SA1294"
```

```
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42554.1"
                     /db_xref="GI:13701260"
                     /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                     HFKIFVPDSPPAILYIVISISLLLIQKQNSIIBALAFVTLFKYGIWAVIMNILFIIEQ
                     GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                     DFITHVWQIGVLSCCLSVFGLLLYIELNKLLKCK"
     gene            complement(1578..2153)
                     /gene="SA1295"
     CDS             complement(1578..2153)
                     /gene="SA1295"
                     /note="ORFID:SA1295"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42555.1"
                     /db_xref="GI:13701261"
                     /translation="MSETLNQIKESFIEYLLFQYRPKSRIAVWVLNYIKVNEAKLANI
                     HFVDTKINHHETLEIAEVGSHASAIQFTKRNIKLMVTNEIFDYIANHNCAFDIQIHFA
                     NVSKREQRLDDLIVAQLTESPSYQTYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                     EKQRFYQLTQILNTLKLVNKHNQFEDLADDD"
     gene            complement(2167..3411)
                     /gene="SA1296"
     CDS             complement(2167..3411)
                     /gene="SA1296"
                     /note="ORFID:SA1296"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42556.1"
                     /db_xref="GI:13701262"
                     /translation="MEDIYKLIDDINLQKLENLDSRVNRAITTDNDDALFILGETLYN
                     FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                     ADLYQQINMMEVAIDKLQRALELEPNDPIIHFALAEMLYDGQYLRATSEYETVLETG
                     EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISVDKNDITQE
                     AIKIMTLLSKDPDYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTGC
                     LRIEHGDANEGVLLLKQALEVDNAYQEPLLILSDLYRNEDYEAIIELLTYVDEEDLD
                     PTFMWHLAFAYGQEERDKEAQHFPELAYPTMKTNIDFMSDYFFVLIEIGQKEQAITIL
                     NQLLELEPSNENWHDELSRLQS"
     gene            complement(3418..4716)
                     /gene="aroA"
     CDS             complement(3418..4716)
                     /gene="aroA"
                     /note="ORFID:SA1297"
                     /codon_start=1
                     /transl_table=11
                     /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                     /protein_id="BAB42557.1"
                     /db_xref="GI:13701263"
                     /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASIABGTSNIYK
                     PLLGEDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAFKTPHQVLYTGNSGTTRLLA
                     GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLLIKPSVIKGIN
                     YQMEVASAQVKSAILFASLFSNDTTVIKELDVSRNHTETMPRHFNIPIEAERLSITTT
                     PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                     GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                     IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                     MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene            complement(4726..5790)
                     /gene="aroB"
     CDS             complement(4726..5790)
                     /gene="aroB"
                     /note="ORFID:SA1298"
                     /codon_start=1
                     /transl_table=11
                     /product="3-dehydroquinate synthase"
                     /protein_id="BAB42558.1"
                     /db_xref="GI:13701264"
                     /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                     YFADKFDDILSYENVHKVIIPAGEKTFEQYQETLEYILSHVTRNTAIIAVGGGAT
                     GDFAGFVAATLLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                     LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEQHFKDREILQSLNGMDKYIAKGI
```

```
ETKLDIVVADEKEQGVRKFPLNLGHTFGHAVEYNHKIAHGHAVMLGIIYQPIVANILFN
SNHDIQHYINYLIKLGYPLETITDIDFETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
QHIDQTTIQHACEQLKTYFK"
```

gene   complement(5816. .6982)
          /gene="aroC"
CDS   complement(5816. .6982)
          /gene="aroC"
          /note="ORFID:SA1299"
          /codon_start=1
          /transl_table=11
          /product="chorismate synthase"
          /protein_id="BAB42559.1"
          /db_xref="GI:13701265"
          /translation="MRYLTSGESHGPQLTIVEGIPANLEIKVEDINKEMFKRQGYG
          RGRRMQIEKDTVEIVSGVRNGYTLGSPIIMVVTNDDFTHWRKIMGAAPISEBERENMK
          RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETARRVAVGALCKVLLQQLDIDIY
          SRVEIGIKDKDFYDSETFKANLDRNDVRVIDDSIAQAMRDKIDEAKNEGDSIGGVV
          QVVVENMVGVGSYVHYDRKLDGKIAQGVVSINAFKGYSFGEGFKAAEKPGSEIQDEI
          LYNSEIGYYRGSNHLGLEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
          IERSDSCAVPAASIVCEHVVAFEIAKALLEEFQSNHIEQLQQQIADRRQLNVEF"

gene   complement(7368. .7568)
          /gene="SA1300"
CDS   complement(7368. .7568)
          /gene="SA1300"
          /note="ORFID:SA1300"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB42560.1"
          /db_xref="GI:13701266"
          /translation="MKNIISILIGIIMFLKLMELLYGAIFLDKPLNPITKIIFILTLI
          YIFYVLVKELIIFLKSKYNKSA"

gene   complement(7597. .7722)
          /gene="SAS045"
CDS   complement(7597. .7722)
          /gene="SAS045"
          /note="ORFID:SAS045"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB42561.1"
          /db_xref="GI:13701267"
          /translation="MWIVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"

gene   complement(7777. .8226)
          /gene="ndk"
CDS   complement(7777. .8226)
          /gene="ndk"
          /note="ORFID:SA1301"
          /codon_start=1
          /transl_table=11
          /product="nucleoside diphosphate kinase"
          /protein_id="BAB42562.1"
          /db_xref="GI:13701268"
          /translation="MERTFLMIKPDAVQRNLIGEVISRIERKGLKLVGGKLMQVPMEL
          AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGEDAVNVSRHIIGSTNPSEASPGSIR
          GDLGLTVGRNIIHGSDSLESAERINLWFNENEITSYASPRDAWLYE"

gene   complement(8318. .9277)
          /gene="gerCC"
CDS   complement(8318. .9277)
          /gene="gerCC"
          /note="ORFID:SA1302"
          /codon_start=1
          /transl_table=11
          /product="heptaprenyl diphosphate syntase component II"
          /protein_id="BAB42563.1"
          /db_xref="GI:13701269"
          /translation="MAKLNMNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
          PAFVILSSQFGKDEQTSEQYQVAVALELIHMATLVHDDVIDKSDKRRGKLTISKKWD
```

```
Qy       1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  291118 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC  291059

Qy      61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAACTACGTATTACAGAA    120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  291058 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAACTACGTATTACAGAA   290999

Qy     121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290998 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA  290939

Qy     181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290938 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA  290879

Qy     241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA   300
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  290878 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA  290819

Qy     301 TATACTTCAGTAGTCTGGTGCACATGGTGCACAGGTTATTATCA                 360
           ||||||||||||||||||||||||||||||||||||||||||||
Db  290818 TATACTTCAGTAGTCTGGTGCACATGGTGCACAGGTTATTATCA                290759

Qy     361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290758 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA  290699

Qy     421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTTAAGT   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290698 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTTAAGT  290639

Qy     481 TATAAAACCTGATTACGCAATTATGACAAATATTGATTCCGATCATCCTGATTATTTAA   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290638 TATAAAACCTGATTACGCAATTATGACAAATATTGATTCCGATCATCCTGATTATTTAA   290579

Qy     541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290578 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT  290519

Qy     601 ATTGCTTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATAT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290518 ATTGCTTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATAT  290459

Qy     661 TATGGGATTTAAAGATTCGGATGACATTTATGCTCAAAAATATTCAAATTACGGATAAAGGT   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290458 TATGGGATTTAAAGATTCGGATGACATTTATGCTCAAAAATATTCAAATTACGGATAAAGGT  290399

Qy     721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATGATCACTTCCTGTCTCCACAATAT   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290398 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATGATCACTTCCTGTCTCCACAATAT  290339

Qy     781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTTATTAGAGAAGCTA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290338 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTTATTAGAGAAGCTA  290279

Qy     841 GATGTTACAAATATTAAAGAAGAGCATTAGAAACGTTGGTGGTTGTTAAACGTCGTTTCAAT   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290278 GATGTTACAAATATTAAAGAAGAGCATTAGAAACGTTGGTGGTTGTTAAACGTCGTTTCAAT  290219

Qy     901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCCAAGAGAAATT   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290218 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCCAAGAGAAATT  290159

Qy     961 AGTGTGCTACAAAGCTGTTGAACAGCAGCGAAAGAAATTCCACATAAAGAAGTTGTTGCAGTATTT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290158 AGTGTGCTACAAAGCTGTTGAACAGCAGCGAAAGAAATTCCACATAAAGAAGTTGTTGCAGTATTT  290099

Qy    1021 CAACCACCACACTTTCTTCTAGAACCACACAGGCATTTTAAATGAATGATTTTGAGAAAGTTTAAGT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290098 CAACCACCACACTTTCTTCTAGAACCACACAGGCATTTTAAATGAATGATTTTGAGAAAGTTTAAGT  290039

Qy    1081 AAAGCAGATGCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGGCGCA  1140
```

```
Db   290038  AAAGCAGAGATCGTGTATTCCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGCGGCA 289979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1141  TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT      1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289978  TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT      289919

Qy     1201  ATTAATGTATTAGAACAATTGATATGCTTATTTATTTATTATGGGTGCAGGTGATATT      1260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289918  ATTAATGTATTAGAACAATTGATATGCTTATTTATTTATTATGGGTGCAGGTGATATT      289859

Qy     1261  CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTAA          1314
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289858  CAAAAATTACAAAATGCATATTTAGATAAATTAGGCCATGCATGAAAAATGCGTTTAA    289805
```

RESULT 6
AP003363/c

LOCUS   AP003363   342600 bp   DNA   linear   BCT 24-APR-2003
DEFINITION   Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
             sequence, section 6/9.
ACCESSION   AP003363 BA000017
VERSION   AP003363.2 GI:14247399
KEYWORDS   .
SOURCE   Staphylococcus aureus subsp. aureus Mu50
  ORGANISM   Staphylococcus aureus subsp. aureus Mu50
             Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
             Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
             Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
             Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
             Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
             Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
             Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE   Whole genome sequencing of meticillin-resistant Staphylococcus
             aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
   MEDLINE   21311952
    PUBMED   11418146
REFERENCE   2 (bases 1 to 342600)
  AUTHORS   Ohta,T.
  TITLE   Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
             of Medical Technology and Nursing, Department of Medical
             Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
             (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
             Fax:81-298-53-3454)
COMMENT   On May 29, 2001 this sequence version replaced gi:13875943.
FEATURES   Location/Qualifiers
     source   1..342600
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /mol_type="genomic DNA"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene   complement(42..815)
                     /gene="SAV1628"
     CDS   complement(42..815)
                     /gene="SAV1628"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB57790.1"
                     /db_xref="GI:14247400"
                     /translation="MKHQFSRNELAIGGQGLDLLKQQTVVILGVGVGVGSFAABZALART
                     NIGHIILIDKDVDITNVNRQLHALTTVGQSKVSLMEERIKLINPDCKVTPLHMFYT
                     EETYEBIFDNDYIDYFVDASDTIMYKVHLMKECLDRGIKVISSMGAANKTDPTRFTIS
                     DISKTHTDPMAKIIRQLKKLGIRKGIPVVFSDESPIVIREDVDIVGDKNAINRKGQ
                     MPPSSNAFVPSTVGLICASYVVNDVLKDIPVRRIKDKGQ"
     gene   complement(1029..1130)
                     /gene="SAV1629"
     CDS   complement(1029..1130)
```

```
                     /gene="SAV1629"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57791.1"
                     /db_xref="GI:14247401"
                     /translation="MGFLAAYTPYMKDVQNKKQGTHLYITGRMIKLL"
     gene            complement(1276..3042)
                     /gene="aspS"
     CDS             complement(1276..3042)
                     /gene="aspS"
                     /note="SAV1630"
                     /codon_start=1
                     /transl_table=11
                     /product="aspartyl-tRNA synthetase"
                     /protein_id="BAB57792.1"
                     /db_xref="GI:14247402"
                     /translation="MSKRTTYCGLVTEAFLGQEITLKGWVNNRRDLGGLIFVDLRDRE
                     GIVQVVFNPAFSEEALKIAETVRSEYVVEVQGTVTKRDPETVNPKIKTGQEVEQVTNI
                     KVINKSETPPFSINEENVNDENIRLKYRILDLRRQELAQTFKMRHQITRSIRQYLDD
                     EGFFDIETPVLTKSTPEGARDYLVPSRVHDGEYALPQSPQLFKQLLMISCGHKYYOI
                     VKCFRDEDLRADRQPEFTQVDIEMSFVDQEDVMQMGEEMLKKVVKEKVGVEINGAFPR
                     MTYKEAMRRYGSDKPDTRFEMELIDVSQLGRDMDFKVFKDTVENDGBIKAIVAKGAAE
                     QYTRKDMDALTEFVNIYGAKGLAWVKVVEDGLTGPIGRPFETENVETLLTLGAEAGD
                     LVMFVADKPNVVAQSLGALRVKLAKELGLIDETKLNFLWVTDWPLLEYDEDAKRYVAA
                     HHPFTSPKEADIAKLGTAPEEABANAYDIVLNGYELGGGSIRIHDGEBLQEKMFEVIGF
                     TKEQAQEQFGFLLDAFKYGAPPHGGIALGLDRLVMLLTRNTNLRDTIAFPKTASAATCL
                     LTNAPGEVSDKQLEELSLRIRH"
     gene            complement(3058..4320)
                     /gene="hisS"
     CDS             complement(3058..4320)
                     /gene="hisS"
                     /note="SAV1631"
                     /codon_start=1
                     /transl_table=11
                     /product="histidyl-tRNA synthetase"
                     /protein_id="BAB57793.1"
                     /db_xref="GI:14247403"
                     /translation="MIKIPRGTQDILPEDSKKWRYIENQLDELMTFYNYKEIRTPIFE
                     STDLFARGVGDSTDVVQKEMYTFKDKGDRSITLRPEGTAAVVRSYIEHKMQGNPNQPI
                     KLYYNGPMFRYERKQGRYRQFNOFGVRAIGAENPSVDAEVLAMVMHIYQSFGLKHLX
                     LVINSVGMASRKYNEALVKHFEPVIHEFCSDCQSRLHTNPMRILDCKVDRDKEAIK
                     TAPRITDFLNEESKAYYEQVKAYLDDLGIPYIEDNLVRGLDYYTHTAFELMMDNPNY
                     DGAITTLCGGGRYNGLLELLDGPSETGIGFALSIERLLLALEEEGIELDIBENLDLFI
                     VTMGDQADRYAVKLLNHLRHNGIKADKDYLQRKIKGQMKQADRLGAKFTIVIGDQELE
                     NNKIDVKNMTTGBSETIELDALVEYFKK"
     gene            complement(4782..5657)
                     /gene="lytH"
     CDS             complement(4782..5657)
                     /gene="lytH"
                     /note="SAV1632"
                     /codon_start=1
                     /transl_table=11
                     /product="N-acetylmuramoyl-L-alanine amidase"
                     /protein_id="BAB57794.1"
                     /db_xref="GI:14247404"
                     /translation="MKKIEAWLSKKGLKNKRTLIVVIAFVLFIFLFLLLNSNSEDSG
                     NITITENAELRTGPNAAYPVIYKVEKGDHFKKIGKVGKWIEVEDTSSNEKGWIAGWHT
                     NLDIVADNTKEKNPLQGKTIVLDPGHGSDQGASSNTKYKSLEKDYTLKTAKELQRTL
                     EKEGAIYKMTRTDTTYVSLENRDIKGDAYLSIHNDALSSSNANGMTVVWYWHDNQRALA
                     DTLDATIQKKGLLSNRGSRQENYVLRQTKVPAVLLELGYISNPTDETMIKDQLHRQI
                     LEQAIVDGLKIYFSA"
     gene            complement(5654..6106)
                     /gene="SAV1633"
     CDS             complement(5654..6106)
                     /gene="SAV1633"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB57795.1"
                     /db_xref="GI:14247405"
                     /translation="MKVVVQRVKEASVTNDTLNNQIKKGYCLLVGIGQNSTEQDADVI
                     AKKIANARLFEDNNKLNFNIQQMNGEILSVSQFTLYADVKKGNRPGFSNSKNPDQAV"
```

```
Query Match                      99.5%;   Score 1310.8;  DB 1;   Length 342600;
Best Local Similarity            99.8%;   Pred. No. 3.3e-190;
Matches 1312;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy      1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132799 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 132740

Qy     61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132739 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 132680

Qy    121 GTTGCTCTTAGAAATAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132679 GTTGCTCTTAGAAATAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 132620

Qy    181 GATATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAATAGTACGTGCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132619 GATATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAATAGTACGTGCA 132560

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132559 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 132500

Qy    301 TATACTTCAGTAGCTGTAACTCGTGCACATGGTAAAACTTCTACAACACAGGTTATTATCA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132499 TATACTTCAGTAGCTGTAACTCGTGCACATGGTAAAACTTCTACAACACAGGTTATTATCA 132440

Qy    361 CATGTTATGAATGGTGATAAAAAGACTCATTTTAATTGGTGATGGCACAGGTATGGGA  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132439 CATGTTATGAATGGTGATAAAAAGACTCATTTTAATTGGTGATGGCACAGGTATGGGA 132380

Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132379 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 132320

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132319 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 132260

Qy    541 GATATTAATGATGTTTTGATGATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132259 GATATTAATGATGTTTTGATGATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 132200

Qy    601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTTCCAATTTATTAT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132199 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTTCCAATTTATTAT 132140

Qy    661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132139 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 132080

Qy    721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCCACAATAT  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132079 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCCACAATAT 132020

Qy    781 GGTGACCATACAGTTTTAAATGCTGTAATTGCGATTAGTTATTTAGGACAGAAGCTA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132019 GGTGACCATACAGTTTTAAATGCTGTAATTGCGATTAGTTATTTAGGACAGAAGCTA 131960

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131959 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 131900

Qy    901 GAAACTACAAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAGAAATT  960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131899 GAAACTACAAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAGAAATT 131840

Qy    961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131839 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 131780

Qy   1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTAAGT 1080
```

gene　complement (6118..8307)
　　　/gene="relA"

CDS　complement (6118..8307)
　　　/gene="relA"
　　　/note="SAV1634"
　　　/codon_start=1
　　　/transl_table=11
　　　/product="GTP pyrophosphokinase"
　　　/protein_id="BAB57796.1"
　　　/db_xref="GI:14247406"
　　　/translation="MNNEYPYSADEVLHKAKSYLSADEYBYVLKSYHIAYEAHKGQFR
KNGLPYIMHPIQVAGILTEMRLDGPTIVAGFLHDVIEDTPYTFEDVKEMFNEEVARIV
DGVTKLKKVKYRSKEFQQAENRKLFIAIAKDVRILVKLADRLHNMRTLKAMPREKQ
IRISRETLEIYAPLAHHLGINTIKWELEDTALRYIDNVQYFRIVNLMKKKRSEREAYI
ETAIDRIRTEMDRMNIEGDINGRPKHIYSIYRKMMKQKKQFDQIFDLLAIRVIVNSIN
DCYAILGLVHTLWKPMGRFKDYIAMPKQNLYQSLIHTTVVGPNGDPLEIQIRTFDMHE
IAEHGVAAHWAYKEGKKVSEKDQTYQNKLNWLKELABADHTSSDAQEFMETLKYDLQS
DKVYAFTPASDVIELPYGAVPIDFAYAIHSVGNKMIGAKVNGKIVPIDYLLQTGDIV
EIRTSKHSYGPSRDWLKIVKSSAKGKISFFKKQDRSSNIEKGRMMVEVELKEQGFR
VEDILTEKNIQVVNEKYNFANBDDLPAAVGFGGVTSLQIVNKLTERQRILDKQRALNE
EQEVTKSLPIKDNIITDSGVVEGLENVLIKLSKCCNPIPGDDIVGYITKGHGIKVHR
TDCPNIKNETERLINVEWVSKDATOKYQVDLEVTAYDRNGLLNEVLQAVSSTAGNLI
KVSGRSDIDKNAIINISVMVKNVNDVTRVVEKIKQLGDVYTVTRVWN"

gene　complement (8735..9253)
　　　/gene="apt"

CDS　complement (8735..9253)
　　　/gene="apt"
　　　/note="SAV1635"
　　　/codon_start=1
　　　/transl_table=11
　　　/product="adenine phosphoribosyl transferase"
　　　/protein_id="BAB57797.1"
　　　/db_xref="GI:14247407"
　　　/translation="MDLKQYVSEVQDMPKPGVSFKDITIMDNGRAYGYATDKIVEYA
KDRDVDIVVGPEARGFIIGCPVAYSMGIGFAPVRKEGKLPREVIRYEYDLEYGTNVLT
MHKDAIKPGQRVLITDDLLATGGTIEAAIKLVEKLGGIVVGIAFIIELKYLNGIEKIK
DYDVMSLISYDE"

gene　complement (9275..11548)
　　　/gene="SAV1636"

CDS　complement (9275..11548)
　　　/gene="SAV1636"
　　　/note="similar to single-strand DNA-specific exonuclease"
　　　/codon_start=1
　　　/transl_table=11
　　　/product="hypothetical protein"
　　　/protein_id="BAB57798.1"
　　　/db_xref="GI:14247408"
　　　/translation="MIKPKYKWKLTKPAEYISDELTSKLKLTPIVKKILESKSIIDEQ
AIESIISDTDVNHDALQLSDMTKAIERIKRAIANDEKILVGDYDADGVTSTTILVTT
LQLLGAQVGWHIPNRFTEGYGHELAPNAHDEGITLIITVDNGIQHNEIKMVQDLLG
VDVIVTDHHEIGSTLPEAYAIVHPMIPSFNYPFQOLGCGAGVAYKLAQALIENVPDYFK
ALVAIGTIADLVSLIDENRSLVKQGLKVLNDQCPTSVKALLKEAGYNDNIDEETIGFI
IGPRLNAVGRIDDASLACELLMTDDEEAAFLAEQVEHFNRERKDIVATITEEAMAMA
ETKVKKGDLPILLAKENWHRGVLGIVASKIVETFALPTLINLDREQNHAKGSARSID
QVSMFEILSAHQELIAKFPGGHHMAAGMTMDIENIESLAEGLNKWMKELSKTTSLDPVK
PVDVLLTENDTTIKNIRDMNRLRPFGTDFSRPIFEMDDLSVSSVKAIGQQKNHLKLTL
GESNIAALFWONGHLEPELQDEQPINILGSVQINEWNGNQSPQLIIQDIAMNEQQILD
YRSKRKSLPFTENDENIVVLIHPKSDKVNANEYYYGEIKQQTDKVVIRDLPTSMEDL
SNSLQQLQFSQLYIVLQHNHSIYFDGIPMNEFFKKCYALITKQNETNIQEGMLLCQH
LSVKPDTLKFMLKVFLDLKFVTQEDGLIRINQQPDKRSIDSSKVYQLRQQRMDVEKQL
LYQDFSEIKNWIKSQUS"

gene　complement (11751..14030)
　　　/gene="secF"

CDS　complement (11751..14030)
　　　/gene="secF"
　　　/note="SAV1637"
　　　/codon_start=1
　　　/transl_table=11
　　　/product="protein-export membrane protein SecDF"
　　　/protein_id="BAB57799.1"
　　　/db_xref="GI:14247409"

KIYEYFNDALRAYGLTVKTGEFGTHMNVSINNDGPVTIIYESQDGKIQ"

```
Db    131779 CACCACACTTTCTCTAGAACCAAGCATTTTAAATGAATTTGCAGAGAAGTTTAAGT 131720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131719 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 131660
Qy      1141 TTAACGAGATACAAGATTAATTAATTGAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131659 TTAACGAGATACAAGATTAATTAATTGAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 131600
Qy      1201 ATTAATGTTATTAGACAATTGATAATGCGTGTTATTTTATTTATGGTGCAGGTGATATT 1260
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131599 ATTAATGTTATTAGAACAATTGATAATGCGTGTTGTTTTATTTATGGTGCAGGTGATATT 131540
Qy      1261 CAAAAATTACAAAATGCATATTTAGATAATAATTAGGCATGAAAAATGCGTTTAA 1314
             ||||||||||||||||||||||||||||||||||||||||||
Db    131539 CAAAAATTACAAAATGCATATTTAGATAATAATTAGGCATGAAAAATGCGTTTAA 131486
```

```
RESULT 7
AP004828/c
LOCUS       AP004828     272850 bp    DNA     linear    BCT 20-DEC-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 7/10.
ACCESSION   AP004828 BA000033
VERSION     AP004828.1  GI:21204850
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
   MEDLINE  22040717
    PUBMED  12044378
REFERENCE   2  (bases 1 to 272850)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
            Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1..272850
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /mol_type="genomic DNA"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(469..1833)
                     /gene="MW1681"
     CDS             complement(469..1833)
                     /gene="MW1681"
                     /note="ORFID:MW1681"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB95546.1"
                     /db_xref="GI:21204851"
                     /translation="MTQQQNNKRTLKNKHTYQNEPLPNRKDFVSFITGALVGSALGL
                     YFKNKYQKADDLKVKEQELSQKFEERKTQLEETVAPTKERVEGFLNKSKNEQAALKA
                     QQAAIKEEVSANNLSDTSQEAQEIQEAKKEAQAETDKSAAVSNEESKAAALKAQQAAI
                     KEEASANNLSDTSQEAQEIQEAKKEAQAETDKSAAVSNEESKAAALKAQQAAIKEEAS
                     ANNLSDTSQEAQEIQEAKREAQAETDKSVAVSNEESKAVALKVSKEESQAERLANAAK
                     QKQAKLTPGSKESQLTEALFAEKPVAKNDLKEIPQLVTKKQNDVSESETVNIDNKDTVK
                     QKEAKFENGVITRKADEKTTNNTAVDKKSGKQSKKTTPSNKRNASKASTNKTSGQKKQ
```

```
                     HNKKSSQGAKKQSSSSKSTQKNNQTSNKNSKTTNAKSSNASKTPNAKVEKANSKIEKR
                     TFND"
     gene            complement(1907..2398)
                     /gene="MW1682"
     CDS             complement(1907..2398)
                     /gene="MW1682"
                     /note="ORFID:MW1682
                     hypothetical protein, similar to general stress protein
                     homolog"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95547.1"
                     /db_xref="GI:21204852"
                     /translation="MDWTLPIAGIIAAIAPLILCIGIVAVLNSVKKNLDVVAKTLDGV
                     BGQVQGITRETTDLLHKVNRLTEDIQGKDRLNSVVDAVKGIGDSVQTLNSSVDRVTN
                     SITHNISQNEDKISQVVQWSNVAMEIADKWQNRHYRRGSANYKANNVATDANHSYTSR
                     VDK"
     gene            complement(2472..3785)
                     /gene="murC"
     CDS             complement(2472..3785)
                     /gene="murC"
                     /note="ORFID:MW1683"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-Aacerylmuramate-alanine ligase"
                     /protein_id="BAB95548.1"
                     /db_xref="GI:21204853"
                     /translation="MTYHFVGIKGSGMSSLAQIMHDLGHEVQSDIENVFTEVALR
                     NKGIKILPFDANNIKEDMVVIQGNAFASSHEBIVRAHQLKLDVVSYNDFLGQIIDQYT
                     SVAVTGAHGKTSTTGLLSHVMNGDKKTSFLIGDTGMGLPESDYFAFEACEYRRHFLS
                     YKPDYAIMTNIDFDHPDYFKDINDVFDAFQEMAHNVKKGIIAWGDEHLRKIEADYPLI
                     YYYGFKDSDDIYAQNIQITDKGTAFDVYVDGFYDHFLSPOYGDHTVLNALAVIAISY
                     LEKLDVTNIKEALETFGGVKRRFNETTIANQVIVDDYAHHPREISATIBTARKKYPHK
                     EVVAVFQPHTFSRTQAFLNEFABSLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEG
                     ASLINEDSINVLEQFDNAVVLFMGAGDIQKLQNAYLDKLGMKNAF"
     gene            complement(3809..7633)
                     /gene="MW1684"
     CDS             complement(3809..7633)
                     /gene="MW1684"
                     /note="DNA translocase stage III sporulation prot homolog
                     ORFID:MW1684"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95549.1"
                     /db_xref="GI:21204854"
                     /translation="MSWFDKLFGEDNDSNDDLIHRKKRRQESQNIDNHDSLLPQNN
                     DIYSRPRGKFRPMSVAYENENVEQSADTISDEKEQYHRDYRKQSHDSRSQKRHRRRR
                     NQTTEEQNYSEQRGNSKISQQSIKYKDHSHYHTNKPGTYVSAINGIEKETHKSSKTHNM
                     YSNNTNHRAKDSTPDYHKESFKTSEVPSAIFGTMKPKLENGRIPVSKPSEKVESDKQ
                     KYDKVYAKTQFSQNKQLEQEKQNDSVVKQGTASKSSDENVSSTTKSMPNYSKVDNTIK
                     IENIYASQIVEBIRRBRERKVLQKRRFKKALQQKREEHKNEEQDAIQRAIDEMYAKQA
                     ERYVGDSSLNDDSDLTDNSTDASQLHTNGIENETVSNDENKQASIQNEDTNDTHVDES
                     PYNYEEVSLNQVSTTKQLSDDEVTVSNVTSQHQSALQHNVEVNDKDELKNQSRLIADS
                     EEDGATNKERYSGSQIDDABFYELNDTEADEDTSNIEDNTNRNASEMHVDAPKTQEH
                     AVTESQVNNIDKTVDNEIELAPRHKDDQTNLVNSLKTNDVNDGVTVEDSSMNEIEK
                     HNAEITENVQNEAAESEQNVEKTIENVNPKKQTEKVSTLSKRPFNVVMTPSDKKRMM
                     DRKKHSKVNVPBLKPVQSKQAVBSKPASQAAPSSRFDSQESNTNAYKTNNMTSNNVE
                     NNQLIGHAETENDYQNAQQYSEQKPSADSTQTEIFEESQDDNQLEDEQVDQSTSSSVS
                     EVSDITEESEETTHPNNTSGGQDNDDQQKDLQPSFSNQNEDTANENRPRTNQQDVATN
                     QAVQTSKPMIRKGPNIKLPSVSLLEEPQVIEBSDEDWITDKKKBLNDALFYFNVPAEVQ
                     DVTEGPSVTRPELSVEKGVVKSRITALQDDIKMALAAKDIRIEAPIPGTSRVGIEVPN
                     QNPTTVNLRSIIBSPSFKNAESKLTVAMGYRINNEPLLMDIAKTPHALIAGATGSGKS
                     VCINSIIMSLLYKNHPEELRLLLIDPKMVELAPYNGLPHLVAPVITDVKAATQSLKWA
                     VBEMERRYKLFAHYHVRNITAFNKKAPYDERMPKIVIVIDELADLMMMAPVEQSIA
                     RIAQKARACGIHMLVATQRPSVNVITGLIKANIPTRIAPMVSSSVDSRTILDSGGAER
                     LLGYGDMLYLGSGMNKPIRVCQTFVSDDEIDDVVDFIKQQREPDYLFEKBELLKKTQT
                     QSQDELFDDVCAPMVNEGHISTSLIQRHFQIGYNRAARIIDQLEQLGYVSSANGSKPR
                     DVYYTEADLNKE"
     gene            complement(7654..8250)
                     /gene="MW1685"
     CDS             complement(7654..8250)
                     /gene="MW1685"
                     /note="ORFID:MW1685
```

phenylalanyl-tRNA synthetase (beta subunit) homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95550.1"
/db_xref="GI:21204855"
/translation="MNLFYNPKYVGDVAFLQIEPVEGELNYNKKGNVVEITNEGNVVG
YNIFEISKDITIEEKGHIKLLTDELVNVFQKRISEAGFDYKLNADLSPKFVVGYVETKD
KHPDADKLSVLNVNVGNDTLQIVCGAPNVEAGQKVVVAKVGAVMPSGMVIKDAELRGV
ASSGMICSMKELNLPNAPEEKGIMVLNDSYEIGQAFFE"

```
Qy     1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGCATGAGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3785 ATGACACACTATCATTTGTCCGGAATTAAAGGTTCTGCATGAGTTCATTAGCACAAATC 3726

Qy    61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3725 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 3666

Qy   121 GTTGCTCTTAGAAAATAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3665 GTTGCTCTTAGAAAATAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 3606

Qy   181 GATATGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAGAAATAGTACGTGCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3605 GATATGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAGAAATAGTACGTGCA 3546

Qy   241 CATCAATTGAAATTAGATGTCGTAAGTTATAATGATTTTTAGGACAGATTAATGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3545 CATCAATTGAAATTAGATGTCGTAAGTTATAATGATTTTTAGGACAGATTAATGATCAA 3486

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGTTTATTATCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3485 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGTTTATTATCA 3426

Qy   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3425 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 3366

Qy   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3365 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 3306

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3305 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 3246

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATATGTGTAAAAAGGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3245 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATATGTGTAAAAAGGTATT 3186

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3185 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAC 3126

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3125 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 3066

Qy   721 ACTGCTTTTGATGTGTATGTGATGTGGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3065 ACTGCTTTTGATGTGTATGTGATGTGGAGTTTTATGATCACTTCCTGTCTCCACAATAT 3006

Qy   781 GGTGACCATACAGTTTAAATGCGTAATTGCATTAGCTGTAATTGCGATTAGTTATTGAGAAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3005 GGTGACCATACAGTTTAAATGCGTAATTGCATTAGCTGTAATTGCGATTAGTTATTGAGAAGCTA 2946

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2945 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTTCAAT 2886

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2885 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCCAAGAGAAATT 2826

Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
```

gene   complement(8279..9136)
          /gene="MW1686"
CDS   complement(8279..9136)
          /gene="MW1686"
          /note="ORFID:MW1686"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAB95551.1"
          /db_xref="GI:21204856"
          /translation="MNTFQMRDKLKERLSHLDVDFKFNRBEETLRIYRTDNNKGITIK
LNAIVAKYEDKCEKIVDEIVYVDERAIAQMADKTLESISSSOIMPVIRATSFDKKTKQ
GVPFIYDEHTAETAVYYAVDLGKSYRLIDESMLEDLKLTBQQIREMSLFNVRKLSNSY
TTDEVKGNTFYFINSNDGYDASRIINTAFLNEIEAACCQGEMLVAVPHQDVLIIADIRN
KTGYDVMAHLTMEFFTKGLVPITSLSFGYKQGHLEPIFILGKNNKQKRDPNVIQRLEA
NRRKFNKDK"

gene   complement(9236..9547)
          /gene="MW1687"
CDS   complement(9236..9547)
          /gene="MW1687"
          /note="ORFID:MW1687
thioredoxin homolog"
          /codon_start=1
          /transl_table=11
          /protein_id="BAB95552.1"
          /db_xref="GI:21204857"
          /translation="MKQLESEQQFESLKQGATVFEFTAGWCPDCRVIEPDLPELEARY
PMFDFVSVDRDKFMDICIENGIMGIPSFLVYKNGELLGSYIGKERKSIEQIDAFLAQY
V"

gene   complement(9612..10688)
          /gene="MW1688"
CDS   complement(9612..10688)
          /gene="MW1688"
          /note="ORFID:MW1688
endo-1,4-beta-glucanase homolog"
          /codon_start=1
          /transl_table=11
          /protein_id="BAB95553.1"
          /db_xref="GI:21204858"
          /translation="MNINKKVTLQRIQTLTELHGAPGFEEEVKNYMTQQMAPVDEFI
ENRMGGFFGVKKSKNPAKRVMIAAHMDEIGFMITNITKNGMIQFTNLGGVANDIWQG
QRLVIKNRNGDKIIGVVSNIPKGHFRRTGSEGAPEIKDITLDIGAQNEDEVRERGIDIGD
TIVPHTPFTQLSEHRYSAKAWDNRYGCVLAIBILELLKDIELDVDLYVGANVQEEVGL
RGAKASAEMIDPDVAFVDVDCSPANDVKSGSQPLSGELGKGTLRIKDGTMILKPVFRDY
LLKLVEAHDIEHQYYMSPGGTDGGEIHKANIGIPTAVIGVCARYIHSTDSVFDIRDYF
AARSLLSEAICNLDNNQIETLQYK"

gene   10775..11086
          /gene="MW1689"
CDS   10775..11086
          /gene="MW1689"
          /note="ORFID:MW1689"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB95554.1"
          /db_xref="GI:21204859"
          /translation="MTYKLKYIIPTTIAVAIVIISTSIIQYINRKRYNPVKLNEVKS
YFMNVKGSYIVYEFVHPETDKYRLVYQGGITTIKNGQNIHYDFYADAYTGEVINIVE
R"

gene   complement(11211..12053)
          /gene="MW1690"
CDS   complement(11211..12053)
          /gene="MW1690"
          /note="ORFID:MW1690"

Query Match　　　　98.8%;　Score 1302.8;　DB 1;　Length 1314;
Best Local Similarity　99.5%;　Pred. No. 2.8e-188;
Matches 1307;　Mismatches 0;　Conservative 0;　Indels 0;　Gaps 0;

```
Qy        1 ATGACACACTACATCATTTGTCGGAATTAAAGGTTCTCGCATGAGTTCATTAGCACAAATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGACACACTACATCATTTGTCGGAATTAAAGGTTCTCGCATGAGTTCATTAGCACAAATC 60

Qy       61 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAATACTGGCGCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAATACTGGCGCA 120

Qy      121 GTTGCTCTTAGAAATAAGGGATAAAATATTACCATTGATGCTAATACATAAAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GTTGCTCTTAGAAATAAGGGATAAAATATTACCATTGATGCTAATACATAAAAGAA 180

Qy      181 GATATGGTAGTTATACAAGGTAATGCATTCGGCAGTAGCCATGGCCATGAAGTAGTACGTGCA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GATATGGTAGTTATACAAGGTAATGCATTCGGCAGTAGCCATGGCCATGAAGAAAAGCACGTGCA 240

Qy      241 CATCAATTGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CATCAAATGAAATTAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300

Qy      301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTATATCA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTATATCA 360

Qy      361 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy      421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480

Qy      481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540

Qy      541 GATATTAATGATGTTTTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATATTAATGATGTTTTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660

Qy      601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660

Qy      661 TATGGATTTAAAGATTCGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TATGGATTTAAAGATTCGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720

Qy      721 ACTGCTTTGATGTGTATGTGGATGTGGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ACTGCTTTGATGTGTATGTGGATGTGGAGTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy      781 GGTGACCATACAGTTTAAATGCAGTTAGCTGTAATTGCGATTAGTTATTGGCGATAGAAGCTA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GGTGACCATACAGTTTAAATGCAGTTAGCTGTAATTGCGATTAGTTATTGGCGATAGAAGCTA 840

Qy      841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900

Qy      901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Qy      961 AGTGCTACAATTGAAACAGCACGAGAAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 AGTGCTACAATTGAAACAGCACGAGAAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020

Db     2825 AGTGCTACAATTGAAACAGCACGAGAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     1021 CAACCACACACTTTCTCAGAACACAGCCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
Db     2765 CAACCACACACTTTCTCAGAACACAGCCATTTTAAATGAATTGCAGAAAGTTTAAGT 2706

Qy     1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAATACTGGCGCA 1140
Db     2705 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAATACTGGCGCA 2646

Qy     1141 TTAACGATACAAGATTTAATTGATAAAATGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
Db     2645 TTAACGATACAAGATTTAATTGATAAAATGAAGGTGCATCGTTAATTAATGAAGATTCT 2586

Qy     1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTTATTGGGTGCAGGTGATATT 1260
Db     2585 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTTATTGGGTGCAGGTGATATT 2526

Qy     1261 CAAAAATTACAAAATGCATATTTAGATAAATAGGCATGAAAAATGCGTTTTAA 1314
Db     2525 CAAAAATTACAAAATGCATATTTAGATAAATAGGCATGAAAAATGCGTTTTAA 2472
```

RESULT 8
AF034076

LOCUS　　　　　AF034076　　　1314 bp　　DNA　　　linear　　BCT 16-JUL-1999
DEFINITION　　Staphylococcus aureus UDP-N-acetylmuramoyl-L-alanine synthetase
　　　　　　　　(murC) gene, complete cds.
ACCESSION　　　AF034076
VERSION　　　　AF034076.1　　GI:2642658
KEYWORDS　　　.
SOURCE　　　　　Staphylococcus aureus
ORGANISM　　　Staphylococcus aureus
　　　　　　　　Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE　　　1　(bases 1 to 1314)
AUTHORS　　　Lowe,A.M. and Deresiewicz,R.L.
TITLE　　　　Cloning and sequencing of Staphylococcus aureus murC, a gene
　　　　　　　　essential for cell wall biosynthesis
JOURNAL　　　DNA Seq. 10 (1), 19-23 (1999)
MEDLINE　　　20029270
PUBMED　　　　10565540
REFERENCE　　　2　(bases 1 to 1314)
AUTHORS　　　Lowe,A.M. and Deresiewicz,R.L.
TITLE　　　　Direct Submission
JOURNAL　　　Submitted (11-NOV-1997) Channing Laboratory, Brigham and Women's
　　　　　　　　Hospital and Harvard Medical School, 181 Longwood Ave, Boston, MA
　　　　　　　　02115, USA
FEATURES　　　　　　　Location/Qualifiers
source　　　　　1..1314
　　　　　　　　/organism="Staphylococcus aureus"
　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　/db_xref="taxon:1280"
gene　　　　　1..1314
　　　　　　　　/gene="murC"
CDS　　　　　　1..1314
　　　　　　　　/gene="murC"
　　　　　　　　/function="cell wall biosynthesis"
　　　　　　　　/note="MurC; UDP-N-acetylmuramate-alanine ligase"
　　　　　　　　/codon_start=1
　　　　　　　　/transl_table=11
　　　　　　　　/product="UDP-N-acetylmuramoyl-L-alanine synthetase"
　　　　　　　　/protein_id="AAB87090.1"
　　　　　　　　/db_xref="GI:2642659"
　　　　　　　　/translation="MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALR
　　　　　　　　NKGIKILPFDANNIKEDMVVIQGNAFASSHEEKARAHQMKLDVVSYNDFLGQIIDQYT
　　　　　　　　SVATVGAHGKTSTTGLLSHVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLS
　　　　　　　　YKPDYAIMTNIDFDHPDYFKDINDVFDAFQBMAHNVKKGIIAWGDDEHLRKIEADVPI
　　　　　　　　YYYGFKDSDDIYAQNIQTDKGTAFPDVVVDGEFYDHFLSPQYGDHTVLNALAVIAISY
　　　　　　　　LBKLDVTNIKEALETFGVKRRFNETTIANQVIVDDYAHHPREISATIETARKKYPHK
　　　　　　　　EVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEG
　　　　　　　　ASLINEDSINVLEQFDNAVVLFKGAGDIQKLQNAYLDKLGMKNAF"
ORIGIN

```
Qy   1021 CAACCACACACTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAACCACACACTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGCAGATCGTGTATTCTTATGTGAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140

Qy   1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCC 1200

Qy   1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTATTTAAGGGTGCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTATTTAAGGGTGCAGGTGATATT 1260

Qy   1261 CAAAAATTACAAAAATGCATATTTAGATAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||||
Db   1261 CAAAAATTACAAAAATGCATATTTAGATAAAATGCGTTTTAA 1314
```

RESULT 9

```
AX191741
LOCUS       AX191741              1314 bp    DNA     linear     PAT 15-AUG-2001
DEFINITION  Sequence 23 from Patent WO0149775.
ACCESSION   AX191741
VERSION     AX191741.1  GI:15209910
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial/cell division composition and method
  JOURNAL   Patent: WO 0149775-A 23 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="Staphylococcus aureus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1280"
ORIGIN

Query Match     98.8%;  Score 1302.8;  DB 6;  Length 1314;
Best Local Similarity   99.5%;  Pred. No. 2.8e-188;
Matches 1307;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

```
Qy     1 ATGACACTACTATCATTTGTCTGCGAATTAAAGGTTCTGGCATGATGAGTTCATTAGCACAAATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGACACACTATCATTTGTCTGCGAATTAAAGGTTCTGGCATGATGAGTTCATTAGCACAAATC 60

Qy    61 ATGCATGAAGATTAGGACATGAAGGATAAAATATATTACCATTTGATGCTAATAACATAAAGAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATGCATGAAGATTAGGACATGAAGGATAAAATATATTACCATTTGATGCTAATAACATAAAGAA 120

Qy   121 GTTGCTCTTAGAAATAAGGGATAAAATAAGGGGTAATAACAAGGTTATACAAGGTAATAATGCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTTGCTCTTAGAAATAAGGGATAAAATAAGGGGTAATAACAAGGTTATACAAGGTAATAATGCA 180

Qy   181 GATATGGTAGTTATACAAGGTGTGAAGTTGTAAGTTGTAATAATGATTTTTAGGACAGATTATTTAGGACAATTATTGATCAA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATATGGTAGTTATACAAGGTGTGAAGTTGTAAGTTGTAATAATGATTTTTAGGACAGATTATTTAGGACAATTATTGATCAA 240

Qy   241 CATCAATTGAAATTAGATGTGTAAGTGTAAGTTATAATGAATTTTTAGGACACAGATTATTGATCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCAATTGAAATTAGATGTGTAAGTGTAAGTTATAATGAATTTTTAGGACACAGAGATTATTGATCAA 300

Qy   301 TATACTTCAGTAGCTGTAACTGTGTGCACATGGTGCACATGTGCACAACTTCTACACAGGTTTATTATCA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATACTTCAGTAGCTGTAACTGTGTGCACATGGTGCACATGTGCACAACTTCTACACAGGTTTATTATCA 360

Qy   361 CATGTTGATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   361 CATGTTGATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTTAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTTAAA 540

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATATGTTAAAAAAGGTATT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATATGTTAAAAAAGGTATT 600

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAGAGCAGAGTGTCCAATTTATTATAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAGAGCAGAGTGTCCAATTTATTATAT 660

Qy   661 TATGGATTAAAGATTCCGATGACATTTATGTCTCAAAATATTCAAAATATTCAAATTACGGATAAAGGT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTAAAGATTCCGATGACATTTATGTCTCAAAATATTCAAAATATTCAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy   781 GGTGACCATACAGTTTAAATGCATTTATAAATGCGATTAGTTATTTAGAGAAGCTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGTGACCATACAGTTTAAATGCATTTATAAATGCGATTAGTTATTTAGAGAAGCTA 840

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTTCAAT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTTCAAT 900

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATATGCACACATCCAAGAGAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATATGCACACATCCAAGAGAAATT 960

Qy   961 AGTGCTACAATGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGTGCTACAATTGCAAATTGAAACAGCACGACAGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
```

```
Db   361 CATGTTGATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACTCACTTTTAAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACTCACTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTAAA 540

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAGAGCAGAGTGTCCAATTTATTATAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAGAGCAGAGTGTCCAATTTATTATAT 660

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATATTCAAATTACGGATAAAGGT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATATTCAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy   781 GGTGACCATACAGTTTAAATGCATTTATAAATGCGATTAGTTATTTAGAGAAGCTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGTGACCATACAGTTTAAATGCATTTATAAATGCGATTAGTTATTTAGAGAAGCTA 840

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTCAAT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATGTTACAAATATTAAAGAAGCATTAGAAAGCGTTGGTGGTGTTAAACGTCGTTCAAT 900

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATATGCACACATCAAGAGAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAACTACAATTGCAAATTCAAGTTATTGTAGATGATATGCACACATCAAGAGAAATT 960

Qy   961 AGTGCTACAATTGCAAATTGCAGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGTGCTACAATTGCAAATTACAGCAGCACGACAGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
```

RESULT 10

```
E35613
LOCUS       E35613
DEFINITION  MurC.
ACCESSION   E35613
VERSION     E35613.1  GI:13019102
KEYWORDS    JP 1999225773-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
```

Acc, E3 5613

REFERENCE   1 (bases 1 to 1351)
AUTHORS   Nicola,G.W. and Martin,K.R.B.
TITLE   MurC
JOURNAL   Patent: JP 1999225773-A 1 24-AUG-1999;
             SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT   OS   Unidentified
             PN   JP 1999225773-A/1
             PD   24-AUG-1999
             PF   03-JUL-1999 JP 1998225115
             PR   03-JUL-1997 US   60/052720
             PI   NICOLA G WALLIS,MARTIN K R BURNHAM
             PC   C12N15/09,A61K31/70,A61K38/53,A61K38/55,A61K39/085,A61K39/395,
             PC   A61K39/395,
             PC   A61K48/00,A61K48/00,C07K16/40,C12N1/21,C12P21/02,C12P21/08, PC
             C12Q1/68,
             PC   G01N33/566//(C12N15/09,C12R1:445),(C12N1/21,C12R1:445), PC
             C12N15/00,A61K37/60,
             PC   A61K37/64,(C12N15/00,C12R1:445)
             CC   Strandedness: Double;
             CC   Topology: Linear;
             FH   Key   Location/Qualifiers
             FT   source   1..1351
             FT   /organism='Unidentified'.

FEATURES   Location/Qualifiers
     source   1..1351
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match   98.1%;   Score 1293.2; DB 6; Length 1351;
Best Local Similarity 99.0%; Pred. No. 7.8e-187;
Matches 1301; Conservative   0; Mismatches 13; Indels 0; Gaps   0;

Qy   1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   81

Qy   61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACAACGTATTACCAGAA   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACAACGTATTTACCAGAA   141

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGGTGCTAATAACATAAAAGAA   201

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA   261

Qy   241 CATCAAATTGAAAATTAGATGTGTAAGTTATAATAATGATATTTTTAGGACAGATTATTGATCAA   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CATCAAATTGAAAATTAGATGTGTAAGTTATAATAATGATATTTTTAGGACAGATTATTGATCAA   321

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA   381

Qy   361 CATGTTAATGAAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 CATGTTAATGAAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   441

Qy   421 TTGGCTGAAAGTGAATTATTTCGCTTTGAGGCATGTGAATATAGAGACGTCACTTTTAAGT   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 TTGCCTGAAAGTGAATTATTTCGCTTTGAGGCATGTGAATATAGAGACGTCACTTTTAAGT   501

Qy   481 TATAAACCTGATTACGCAATTATGACAAAATATTCGATCATCCTGATTATTTTAAA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTCAAA   561

Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT   621

Qy   601 ATTGCTTGGGGTGATGATGAACATCTTCACGTAAAAATTGAAGCAGAGATGTTCCAATTATTAT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ATTGCTTGGGGTGATGATGAACATCTTCACGTAAAAATTGAAGCAGAGATGTTCCAATTTATTAC   681

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGTCAAATATTCAAATATTCAAATTACGGATAAAGGT   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TATGGATTAAAGATTCGGATGACATTTATGTCAAATATTCAAATATTCAAATTACGGATAAAGGT   741

Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCATCACTTCCTGTCTCCACAATAT   801

Qy   781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATGCAATTAGTATTATTAGAGAAGCTA   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATGCAATTAGTATTATTAGAGAAGCTA   861

Qy   841 GATGTTACAAATATAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GATGTTACAAATATAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT   921

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT   981

Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAATTGACACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT   1041

Qy 1021 CAACCACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1042 CAACCACCACACTTCTCTAGAACACAAGCATTTTAAATGAATTTGCAGAAAGTTTATGT   1101

Qy 1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCCCA   1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGCTCAATTAGAGAAAATTCTGGGCCA   1161

Qy 1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTCATTAATGAAGATCTT   1221

Qy 1201 ATTAATGTATTAGAACAATTGATAATGTCGTTATTTATTGTTTATTATGTGGTGCAGGTGATATT   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1222 ATTAATGTATTAGAACAATTGATAATGTCGTTGTTTATTTATTATGTGGTGCAGGTGATATT   1281

Qy 1261 CAAAAATTACAAAAATGCATATTTAGATAAAATCGTATTTAGATAAAATCGGTTTTAA   1314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1282 CAAAAATTACAAAATGCATATTTAGATAAAATTAGATAAAATCGGTTTTAA   1335

RESULT 11
BD178030
LOCUS   BD178030   1351 bp   DNA   linear   PAT 16-APR-2003
DEFINITION   MurC.
ACCESSION   BD178030
VERSION   BD178030.1 GI:30015294
KEYWORDS   JP 2002300888-A/1.
SOURCE   Unidentified
  ORGANISM   unidentified
             unclassified.
REFERENCE   1 (bases 1 to 1351)
  AUTHORS   Wallis,N.G. and Burnham,M.K.R.
  TITLE   MurC
  JOURNAL   Patent: JP 2002300888-A 1 15-OCT-2002;
             SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM PLC
COMMENT   OS   Unidentified
             PN   JP 2002300888-A/1
             PD   15-OCT-2002
             PF   25-DEC-2001 JP 2001391079
             PR   03-JUL-1997 US   60/052720
             PI   NICOLA G WALLIS,MARTIN K R BURNHAM
             PC   C12N15/09,A61K31/7088,A61K38/43,A61K38/55,A61K45/00,A61K48/00,
             PC   A61P1/00,

```
PC  A61P7/02,A61P9/00,A61P11/00,A61P13/02,A61P13/12,A61P17/00, PC
    A61P17/02,
PC  A61P19/02,A61P19/08,A61P21/00,A61P25/00,A61P27/00,A61P29/00,
PC  A61P31/04
PC  A61P39/02,A61P43/00,C07K14/195,C12N1/15,C12N1/19,C12N1/21, PC
    C12N5/10,
PC  C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00, PC
    C12N5/00,
PC  A61K37/48,A61K37/64
CC  MurC
FH  Key             Location/Qualifiers
FT  source          1..1351
                    /organism='Unidentified'.

FEATURES            Location/Qualifiers
    source          1..1351
                    /organism="unidentified"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32644"

ORIGIN

Query Match     98.1%;     Score 1293.2; DB 6;   Length 1351;
Best Local Similarity  99.0%;    Pred. No. 7.8e-187;
Matches 1301;  Conservative  0;  Mismatches 13;  Indels  0;  Gaps  0;

Qy        1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGCATGAGTTCATTAGCACAAATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGCATGAGTTCATTAGCACAAATC 81

Qy       61 ATGCATGATTTAGGACATGAAGTTCAAGGATCCGATATTGAGAACTACGTATTACAGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       82 ATGCATGATTTAGGACATGAAGTTCAAGGATCCGATATTGAGAACTACGTATTACAGAA 141

Qy      121 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTGGTGCTAATAACATAAAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      142 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTGGTGCTAATAACATAAAAGAA 201

Qy      181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 261

Qy      241 CATCAAATTGAAATTAGATGTTGTAAGTTATAAATGATTTTTAGGACAGATTATTGATCAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      262 CATCAAATTGAAATTAGATGTTGTAAGTTATATATGATTTTTAGGACAGATTATTGATCAA 321

Qy      301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACACAACGGTTTATTATCA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACACAACGGTTTATTATCA 381

Qy      361 CATGTTATGAATGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      382 CATGTTATGAATGTGATGAAAAGAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy      421 TTGCCTGAAAGTAATATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTTAAGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      442 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTTAAGT 501

Qy      481 TATAAAACCTGATACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTTAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      502 TATAAAACCTGATACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTCAAA 561

Qy      541 GATATTAATGAATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      562 GATATTAATGAATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 621

Qy      601 ATTGCTGGGGTGATGATGAACATCTACGTAAATTGAAGCAGAGTGTTCCAATTTATTAT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      622 ATTGCTGGGGTGATGATGAACATCTACGTAAATTGAAGCAGAGTGTTCCAATTTATTAC 681

Qy      661 TATGGATTTAAAGATTCCGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      682 TATGGATTTAAAGATTCCGATGACATTTATGCTCAAATATTCAAAATTACGGATAAAGGT 741

Qy      721 ACTGCTTTTGATGTGTATGTGGATGTGGAGTTTTATGATGACACTTCCTGTCTCCACAATAT 780
```

```
Db      742 ACTGCTTTTGATGTGTATGTGGATGTGGAGTTTTATGATGATCACTTCCTGTCTCCACAATAT 801
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      781 GGTGACCATACAGTTTTAAATGCATTGGGATTATTGCGATTAGTTATTTAGAGAAGCTA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      802 GGTGACCATACAGTTTTAAATGCATTAGTGGATTAGTTATTTAGAGAAGCTA 861

Qy      841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 921

Qy      901 GAAACTACAAATGCAAATCAAGTTATGTGAGATGATTATGCACACATCCAAGAGAAATT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      922 GAAACTACAAATGCAAATCAAGTTATTGTAGATGATTATGCACACATCCAAGAGAAATT 981

Qy      961 AGTGCTACAAATTGAAACAGCACGAAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      982 AGTGCTACAAATTGACACAGCACGAAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTT 1041

Qy     1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTTAAATGAAATTGCAGAAAGTTTAAGT 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1042 CAACCACACACTTTCTCTAGAACACAAGCCATTTTTAAATGAAATTGCAGAAAGTTTATGT 1101

Qy     1081 AAAGCAGATGCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGCGCA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1102 AAAGCAGATGCGTGTATTCTTATGTGAAATTTTGGCTCAATTAGAGAAAAATTCTGGCGCA 1161

Qy     1141 TTAACGATACAAGATTTAATTGATAAAATGAAGGTGCATCGTAATTAATGAAGATTCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1162 TTAACGATACAAGATTTAATTGATAAATTGGAGGTGCATCGTTCATTAATGAAGATCTT 1221

Qy     1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTTATTATGGGTGCAGGTGATATT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1222 ATTAATGTATTAGAACAATTGATAATGCGTGTTGTTTTATTATGGGTGCAGGTGATATT 1281

Qy     1261 CAAAAATTACAAAATGCATATTTAGATAAAAATGCGTTTTAA 1314
            ||||||||||||||||||||||||||||||||||||||||||
Db     1282 CAAAAATTACAAAATGCATATTTAGATAAAAATGCGTTTTAA 1335
```

```
RESULT 12
AE016748/c
LOCUS       AE016748          300029 bp    DNA       linear    BCT 22-APR-2003
DEFINITION  Staphylococcus epidermidis ATCC 12228, section 5 of 9 of the
            complete genome.
ACCESSION   AE016748 AE015929
VERSION     AE016748.1  GI:27315631
KEYWORDS    .
SOURCE      Staphylococcus epidermidis ATCC 12228
  ORGANISM  Staphylococcus epidermidis ATCC 12228
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 300029)
  AUTHORS   Zhang,Y., Ren,S., Li,H., Fu,G., Lu,L., Lu,G., Jia,J., Tu,Y.,
            Qin,Z., Chen,Z. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) Chinese National Human Genome Center at
            Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES            Location/Qualifiers
    source          1..300029
                    /organism="Staphylococcus epidermidis ATCC 12228"
                    /mol_type="genomic DNA"
                    /strain="ATCC 12228"
                    /db_xref="ATCC:12228"
                    /db_xref="taxon:176280"
    gene            complement(137..1123)
                    /locus_tag="SE1168"
    CDS             complement(137..1123)
                    /locus_tag="SE1168"
                    /codon_start=1
                    /transl_table=11
                    /product="thioredoxin reductase-like protein"
                    /protein_id="AAO04767.1"
                    /db_xref="GI:27315632"
```

Left column

```
                 /translation="MQTIESIIIGGPCGLSAAIEQKKKGIETLVIEKGNVVESIYNY
                 PTHQTFFSSSDKLSIGDIPFIVEDSKPRRNQALVYREVVKGHQLNIHPFEVLITVKK
                 INNKFAITTTKGVTVATGYGQHNTLEAEGAELPKVFHYFKEAHPYFNQNVV
                 IIGGKNSAVDAALELEKAGANVTVLYRGEQYPKAIKPWILPNFESLVNHEKITMEFNA
                 TVTKITDHSVTYEKDGQLIEIDNDYVFAMIGYHPDYDFLKTIGIDIHTNEYGTAPVYN
                 RETPETNVENCYIAGVIAAGNDANTIFIENGKYHGGVITQSILTKQTPLET"
gene             complement(1256..2674)
                 /locus_tag="SE1169"
CDS              complement(1256..2674)
                 /locus_tag="SE1169"
                 /codon_start=1
                 /transl_table=11
                 /product="elastin binding protein"
                 /protein_id="AAO04768.1"
                 /db_xref="GI:27315633"
                 /translation="MIELIKMEGMIVVSNNNFKDDFEKNRQSINPDEQQTELKEDDKT
                 NENKKEADSQNSLSNNSNQQFPPRNAQRRKRRRETATNQSKQQDDKHQKNSDAKTTEG
                 SLDDRYDRAQLQQOHDKSQQONKTEKSQQONRMKDQKDAAIVNGTSESPEHKSKSTQN
                 RPGPKAQQQKRKSESTQSKPSTNKDKKAATGAGIAGAAGVAGAAETSKRHHNKKDKQD
                 SKHSNHENDEKSVKNDDQKQSKKGKKAAVGAGAAAGVGAAGVAHHRNNQNKHHNEEKNS
                 NQNNQYNDQSEGKKKGGFWKILLPLIAAILILGATAIFGGMALNNHNDSKSDDQKIAN
                 QSKKDSDKKDGAQSEDNKDKKSDSDKNADDDSDNSSSNPNATSTNNNDNV
                 ANNNSNYTNQNQQDDNANQNSNNQQATQGGQSHTVFGQENLYRIAIQYYGRGTQANVDK
                 IKRANGLSSNNIHHNGQTLVIPQ"
gene             complement(2795..4174)
                 /locus_tag="SE1170"
CDS              complement(2795..4174)
                 /locus_tag="SE1170"
                 /codon_start=1
                 /transl_table=11
                 /product="probable ATP-dependent DNA helicase RecQ"
                 /protein_id="AAO04769.1"
                 /db_xref="GI:27315634"
                 /translation="MLKDNLKEWFGPDEFKEGQEIIDSVLHGRHTLGILPTGSGKSL
                 CYQLPTYILEKPTLIISPLISLMDDQVMQTKLKGETHVACVHSGMDEVEKRDNIKRLS
                 KSRFIYLSPEYILQPHNFKYIAHINFGLIVLDEAHCLSEWGYDFRPHYALVGKIIHHF
                 NSATVLALTATASPQLETDLSHILSLKLNVVQKSMNRENISLQHYNFNDDEGKIEWLL
                 PFLSNSGPTIIYVSSKKRCLELAQMIYDSGYLTGLYHGDLSYQERQTVQQQFFNNDIP
                 VIVATSAFGMGINKDIRTVIHFLSSSPSNYLQEIGRAGRDGKQSQAISLFQPDDSF
                 ILETLLFTDIITDDDITMETGNHLPDEKKILSTLNIQPTFSQLKDIFHQSYQRRRL
                 GYMRMMGYTNLDQCRRKYLLEFFGEHPQSPKQCCEQDSRLEPIQILNRKKVKRKLSFN
                 EKLYNLFET"
gene             complement(4161..5123)
                 /locus_tag="SE1171"
CDS              complement(4161..5123)
                 /locus_tag="SE1171"
                 /codon_start=1
                 /transl_table=11
                 /product="conserved hypothetical protein"
                 /protein_id="AAO04770.1"
                 /db_xref="GI:27315635"
                 /translation="MYNILSYAMSHAFNYKTNKSIYNILTGKKSHQTFFDASSQQLLS
                 LYHSLPNLKYSTFEQFILQKDDFFKSIQVKIHPQYTYDSLTQTFSCIQLLIQTLSHTR
                 KESNTFIPIVQNYIQQRVKQLYHQVIESNQVSNTIDEIYLLFENLNNKYNHTFLHYY
                 LQGYEESMYTRQQISLIEGIPQSELFEREMENLIDILNQLKDSTKYPILSQAIILSPL
                 LTNTYLSYQKLKSGLNLKEIAQLQNVKLNTIEDHILEMYIKGYLIDYTLFINKKDILE
                 FINYYQKHRGERLKFYKEHFTDWTYFQIKIVIGIERGDLIAER"
gene             5232..5480
                 /locus_tag="SE1172"
CDS              5232..5480
                 /locus_tag="SE1172"
                 /codon_start=1
                 /transl_table=11
                 /product="ferredoxin"
                 /protein_id="AAO04771.1"
                 /db_xref="GI:27315636"
                 /translation="MAKYTIVDMDTCIACGACGAAAPDIYDYDDEGIAYVILDDNQGT
                 AEVPEELEVEDMEDALEGCCPTDSIKIBEDEPPFDGDALKFE"
gene             complement(5572..5676)
                 /locus_tag="SE1173"
CDS              complement(5572..5676)
                 /locus_tag="SE1173"
                 /codon_start=1
```

Right column

```
                 /transl_table=11
                 /product="hypothetical protein"
                 /protein_id="AAO04772.1"
                 /db_xref="GI:27315637"
                 /translation="MSSLVPNKQPQMNHIKQFIWGSIHLIADYFIVYG"
gene             complement(5885..6424)
                 /locus_tag="SE1174"
CDS              complement(5885..6424)
                 /locus_tag="SE1174"
                 /codon_start=1
                 /transl_table=11
                 /product="conserved hypothetical protein"
                 /protein_id="AAO04773.1"
                 /db_xref="GI:27315638"
                 /translation="MQQNKRLITISMLSAVAFVLTFIKFPLPFIPPYLTLDFSDVPTL
                 LATFLLSPIAGIIVALIKNILNFLFNIGDPVGPVANFLAGVSFLLSSYYVRKRKONNR
                 SLIYGLIHGTIVMTIVLSILNYFVLLPLYGMIFNLGDVLNNVKIVIVSGVIPFNLIKG
                 IIISIIFVLLPRRLRHIIK"
gene             complement(7016..8782)
                 /locus_tag="SE1175"
CDS              complement(7016..8782)
                 /locus_tag="SE1175"
                 /codon_start=1
                 /transl_table=11
                 /product="staphylococcal respiratory response protein
                 SrtB"
                 /protein_id="AAO04774.1"
                 /db_xref="GI:27315639"
                 /translation="MNRLNSVVIKLWLTILIVATVLILLSVSLIITFIQYYFTQETEN
                 AIKSDAKRISSLVEKSHNKTLAIQNSQTLDGDPGGLIIMDMSEKPDSSYKDTKKQ
                 MFNEIKKSTKFKKVFKEGEYETQNITIKNKGNSQSYLLLGYPMKAQKGAQSHYSGVFI
                 YKDLKSIEDTNAITIIILITAIFTIASTFAFFLSNRITKPLRQLKTQAQKVSEGD
                 YSQISTVATKDEIGDLSRAFNNMVEIQEHIKAISSKNIRDTLLNSMVEGVLGINNQ
                 REIILSNKMADDIMRHIDDFSKESIEQQIEATFESQQNEYLELEINTRYYVFISSYID
                 RIQTNGRSGIVMVIRDMTNEHNLDQMKDFIANVSHELRTPISLLQGYTESIVDGIVT
                 EPDEIRDSLAIVLDESKRLNRLVNELLNVARMDARGLSVEKELQPIQHLLDRMESKYR
                 MQSRELGLTMTFDSNNDEQLWNYDMDRMDQVLTNLIDNATRYTQAGDSIKISIDEDSD
                 FNILTITDTGTGIAPEHLKQVFDRFYKVDAARKRGKQGTGLGLFICKMIIEEHGGRID
                 VBSELGKGTSFIIRLPKSKQIS"
gene             complement(8766..9491)
                 /locus_tag="SE1176"
CDS              complement(8766..9491)
                 /locus_tag="SE1176"
                 /codon_start=1
                 /transl_table=11
                 /product="staphylococcal respiratory response protein
                 SrA"
                 /protein_id="AAO04775.1"
                 /db_xref="GI:27315640"
                 /translation="MTNEILIVDDEDRIRRLLKILYLERESFEIHRARDGKEAYEALLE
                 NNYACILLDLMLPEMDGIEVASKLREHKDTPIIMLTAKGERTNRVEGFESGADDYIVK
                 PFSPREVVLRVKALLRRTQTANTEQSEPHARDIIEFSHLVIDNDAHRVLADDQQVNLT
                 PKEYELLYLAKTPNKVFDREQLKEVWHYEBYGDLRTVDTHVKRLREKLNRVSSDAA
                 QMIQTVWGVGYKFEVKPSDEPTK"
gene             complement(9625..10362)
                 /locus_tag="SE1177"
CDS              complement(9625..10362)
                 /locus_tag="SE1177"
                 /codon_start=1
                 /transl_table=11
                 /product="ribosomal large subunit pseudouridine synthase
                 B"
                 /protein_id="AAO04776.1"
                 /db_xref="GI:27315641"
                 /translation="MSKEQERLQKRIANSGYTSRRKAEVLITEGRVKVNGQTVTELGS
                 KVMPSDLVEVDGIKIEQEDKIYILFYKPTQVITSVSDDRGRRVVTDYFDDLESRIYPV
                 GRLDYDTSGVLLLTNDGEPTNLMTHPRYHIKKKYVAKLKGYLMREEIKQLEKGIELED
                 GYTQPAQVKIKKQDKEKNTLVEITISEGRNRQVRRMFEYFGHQVNKLTRIQFGPIDL
                 KGLNAGEGRVLTPHEVKILRHLAENGK"
gene             complement(10355..10897)
                 /locus_tag="SE1178"
CDS              complement(10355..10897)
                 /locus_tag="SE1178"
```

```
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO04777.1"
          /db_xref="GI:27315642"
          /translation="MDNIAILEALLYTSGDEGLEQKQIIDILDINLNQLEDLVSKYHS
          HGLITIQRYGSTYVLJTKKETSTYIEQLVKEKSKMKLSQAAMETLSIIAYNQPLTRGDI
          EMIRGINSDGAVKTLIARGLVEAKDVDHSRSHHLITTDLFLNVFGIENLDALPTTEED
          EAEMEDFFSNLVNQKGESNE"
gene      complement (10890..11690)
          /locus_tag="SE1179"
CDS       complement (10890..11690)
          /locus_tag="SE1179"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO04778.1"
          /db_xref="GI:27315643"
          /translation="MYSCYNNFVKYALTMQFDIEVDVMYEVKLDAPNGPLDLLLHLIQ
          KYEIDIYDIPMKALTEQYMQYVHAMNQLEINVASEYLVMASELLMIKSKLLLPQTSIR
          EDIEEDPREDLVGRLIEYQNYKEYTEILNTMKEERDLYFTKHPTDLTHLETNESWDPN
          QTIDLTELIVAYQRVXNRVELNTPKSVEIKKETFTIQQATAQVTERLKQHESFNFFSL
          FTFHEPVEQVVTHFLAILEMSKSGIVNIKQFDIDIIRGVNYSIG"

Query Match      69.7%;  Score 918.8;  DB 1;  Length 300029;
Best Local Similarity  81.2%;  Pred. No. 9.4e-131;
Matches 1067;  Conservative   0;  Mismatches 247;  Indels  0;  Gaps    0;

Qy     1 ATGAGACACTATCATTTTGTCGGAATTAAAGGTTCTGCATGAGTTCATTAGCACAAATC 60
         |||||||||||||||||||||||||| |||||| ||||||||||| |||| |||||||||
Db 248525 ATGAGACACTATCATTTTGTCGGAATTAAAGGCTCAGGCATGAGTTCATTAGCACAAATC 248466

Qy    61 ATGCATGATTTAGGACATGAAGTTCAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
         ||||||| |||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db 248465 ATGCATGACCTCGGTCATGAAGTACAAGGTTCAGACATAGAATCATACGTTTTTACAGAA 248406

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGGTAATAACATAAAAGAA 180
         |||||| ||||||||||||||||||||||||| |||||| ||||||||  |||||||||||
Db 248405 GTTGCATTAGAAATAAAGGGATTAAAAATTTACCTTTTTGATGGCAAATAATATTACAAAA 248346

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
         ||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 248345 GAAATGGTTGTCATCCAAGGTAATGCATTCCCTGATAATCATGAAGAAATTGTTAGGGCA 248286

Qy   241 CATCAATTGAAATTAGATAGTGTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA 300
         ||||||||||||| ||||||||||||||||| || ||||||| |||||||||||||||||
Db 248285 CATGAATTAAAGCTTGATATTATAAAATATCATGACTTTCTTGGTCATGTTATAAATCAA 248226

Qy   301 TATACTTCAGTAGCTGTAACTGGTCACATGGTAAAACTTCTACAACAGTTTATTATCA 360
         |||| ||||||||||||| |||||||||| ||||||||| || ||||| ||||||||||
Db 248225 TATACATCTGTTGCTGTTACAGGTGCACACGCGGTAAAACGTCAACAACTGGTTGTGTCA 248166

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGAA 420
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||| |||
Db 248165 CATGTGATGAATGGAGATAAAAGAGACATCATTCCTAATCGGTGATGGTACAGGTATGGGA 248106

Qy   421 TTGCCTGAAAGTAGTTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
         |||||||||||||||||||||||||| ||||||||||||| ||||||||| || |||| |||
Db 248105 CTACCAGGAAGTGACTATTTTGCTTTTGAGGCTTGTGAATACCGTAGGCACTTCCTTAAGT 248046

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 540
         ||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||||
Db 248045 TATCATCCAGATTATGCAATTATGCATATTGATTTTGACCATCCAGATTATTTTAAA 247986

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
         ||| || ||||||||||||||| |||| |||||||||||||||| || |||||||| |||
Db 247985 AATATCGATGATGTTTATGATGCATTCCAACATATGGCATTAAATGTTAAGAAGGAATT 247926

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAGTAAAAATTGAAGCAGATGTTCCAATTTATTAT 660
         |||||| ||||||||||||||||||||||||||||||||| |||||||||||||| |||||||||||
Db 247925 ATTGCTTGGGGCGGATGATGAATATTTTACGTAAATTAGATGTGGATATTCCGGTTTATTAT 247866

Qy   661 TATGGAGATTAAAGATTCGGATGACGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720
```

```
Db 247865 TATGGCTTTAAAGAAACAGATGACATCTATGCTAAAAATATTCAAAATTACTGGATAAAGGGT 247806
          ||||| ||||||||||||||||||| ||||||| |||||||||||||||||| |||||||||
Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
          |||||||| || ||||||||| |||||||| ||| ||||| ||||| ||| |||||||||
Db 247805 ACGCAATTTGATGTATATATTAAAGGGCAATTTATGATGATCAATTCTTATCCCCACAATAT 247746

Qy   781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGTTAGTTATTAGAGAAGCTA 840
          |||||||||||| |||||||||||| |||||||||| |||||||| |||||| |||||| |||
Db 247745 GGAAATCACACATTCTAAACGCACTCGCTGTGATAATAGCAATTAGCAATTAGTTATTAGAAGACATG 247686

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACAGTTTGTGGTGTGAAACGTCGTTTCAAT 900
          |||||| |||||||||||||||| ||||| ||||||||||| |||||||| ||||| |||
Db 247685 AATGTTAGAAAATATTAAAGAAGCACTCGATAACATTTGGTGGTGTAAAACGTCGTTTTAAT 247626

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTGAGATGGTATTGTGAGATGATGTATGCACACCATCCCAAGAGAAATT 960
          |||||||||| |||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db 247625 GAAACAAAAGTATCAATCAAGTAATACTGATAGTAGACGATTATGCCACCATCCAACAGAGAAATT 247566

Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAGAAGTTGTTGCAGTATT 1020
          ||||||||||||||| |||||||||| || ||||||||||| || || ||| ||||| |||
Db 247565 AGTGCTACAATTGAGACAGCAGCGCGAAAAATATCCGCAAAAGATGTTGTTGCCGTTTTT 247506

Qy  1021 CAACCACACTTCTCTAGACACACAGCATTTTAAATGAATTGCAGAAGTTAAGT 1080
          ||||||||| |||||||||||| || ||||||||||| |||||| || || |||
Db 247505 CAACCACATACGTTCTCAAGAACTCAAGACATTTTAAACGAATTGCTGAGAGTTAAGT 247446

Qy  1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          ||||||||| |||||| ||| ||||||| |||| || |||||||||||||||||||||||
Db 247445 AAGGCAGACCAGTATTTTATGTGAAATATTCGGTTCAATAGAGAAAATACGGGAGAT 247386

Qy  1141 TTAACGATACAAGATTTAATTGATAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          |||| |||||||||||||| ||| |||||||||||| || || |||||||||| ||||||
Db 247385 TTAACTATAGAAGATTTAATTGATAAGATTTGATACCGCTGTAATTTTATTATTATGGGTTTATTTATGGGG 247326

Qy  1201 ATTAATGTATTAGAACAATTTGATAAGTCGTTATTTGATTTATTTATGGTGCAGGTGATATT 1260
          ||||||||||||||||||||| ||||||||| ||||||||||||| |||||||| |||||||
Db 247325 ATTGATGTATTAGAAAATTTGATAACCGCTGTAATTTTATTTATTTATGGGGCTGTGACATT 247266

Qy  1261 CAAAAATTACAAAATGCATATTTGAATAAATGTGAAAAATGCGTTTTAA 1314
          ||||||||||| |||||||||||||| ||||| ||||||| ||| |||||
Db 247265 CAAAGTTACTAAAAGCTTATTTCGAAAAAATTAGGCGTAAAAAAATGATTTTTAA 247212
```

```
RESULT 13
E35614
LOCUS        E35614                     660 bp    DNA     linear     PAT 18-JUN-2001
DEFINITION   MurC.
ACCESSION    E35614
VERSION      E35614.1     GI:13019103
            JP 1999225773-A/2.
KEYWORDS     JP 1999225773-A/2
SOURCE       unidentified
  ORGANISM   unidentified
            unclassified.
REFERENCE    1  (bases 1 to 660)
  AUTHORS    Nicola,G.W. and Martin,K.R.B.
  TITLE      MurC
  JOURNAL    Patent: JP 1999225773-A 2 24-AUG-1999;
            SMITHKLINE BEECHAM CORP, SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT      OS   Unidentified
            PN   JP 1999225773-A/2
            PD   24-AUG-1999
            PF   03-JUL-1998 JP 1998225115
            PR   03-JUL-1997 US     60/052720
            PI   NICOLA G WALLIS,MARTIN K R BURNHAM
            PC   C12N15/09,A61K31/70,A61K38/53,A61K38/55,A61K39/085,A61K39/395,
            PC   A61K39/395,
            PC   A61K48/00,A61K48/00,C07K16/40,C12N1/21,C12P21/02,C12P21/08, PC
            C12Q1/68,
            PC   G01N33/566//(C12N15/09,C12R1:445),(C12N1/21,C12R1:445),
            C12N15/00,A61K37/60,
            PC   A61K37/64,(C12N15/00,C12R1:445)
            CC   Strandedness: Double;
```

```
JOURNAL     Patent: JP 2002300888-A 2 15-OCT-2002;
            SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM PLC
COMMENT     OS   Unidentified
            PN   JP 2002300888-A/2
            PD   15-OCT-2002
            PF   25-DEC-2001 JP 2001391079
            PR   03-JUL-1997 US 60/052720
            PI   NICOLA G WALLIS,MARTIN K R BURNHAM
            PC   C12N15/09,A61K31/7088,A61K38/43,A61K38/55,A61K45/00,A61K48/00,
            PC   A61P1/00,
            PC   A61P7/02,A61P9/00,A61P11/00,A61P13/02,A61P13/12,A61P17/00, PC
            A61P17/02,
            PC   A61P19/02,A61P19/08,A61P21/00,A61P25/00,A61P27/00,A61P29/00,
            PC   A61P31/04,
            PC   A61P39/02,A61P43/00,C07K14/195,C12N1/15,C12N1/19,C12N1/21, PC
            C12N5/10,
            PC   C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00, PC
            C12N5/00,
            PC   A61K37/48,A61K37/64
            CC   MurC
            FH   Key              Location/Qualifiers
            FT   source           1..660
                                  /organism='Unidentified'.

FEATURES             Location/Qualifiers
     source          1..660
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Query Match             48.4%;  Score 637.8;  DB 6;  Length 660;
Best Local Similarity   98.9%;  Pred. No. 2.3e-87;
Matches  642;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy    666 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGC 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGC 60

Qy    726 TTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGGTGA 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGGTGA 120

Qy    786 CCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTGAGAAGCTAGATGT 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTGAGAAGCTAGATGTGTGA 180

Qy    846 TACAAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTCAATGAAAC 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TACAAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTTCAATGAAAC 240

Qy    906 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 300

Qy    966 TACAATTGAAACAGCACGAAGAAATAATATCCACATAAAGAAGTTGGTGTTGCAGTATTCAACC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TACAATTGACACAGCACGAAGAAATAATATCCACATAAAGAAGTTGGTGTTGCAGTATTCAACC 360

Qy   1026 ACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGTAAAGC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGTAAAGC 420

Qy   1086 AGATCGTGTATTCTTATGTGAGAAATTTTGGATCAATTGAGAAATACTGGCGCATTAAC 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGATCGTGTATTCTTATGTGAGAAATTTTGGATCAATTGAGAAATACTGGCGCATTAAC 480

Qy   1146 GATACAAGACATTTAATTAATTGATAAAATTGAAGGTGCATCGTGATAAAATGAAGATTCATTAA 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATACAAGACATTTAATTAATTGATAAAATTGAAGGTGCATCGTGATAAAATGAAGATTCATTAA 540

Qy   1206 TGTATTAGAACAATTTGATAATGCTGTTATTTATTTATGGGTGCAGGTGATATTCAAAA 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TGTATTAGAACAATTTGATAATGCTGTTGTTTTTATTTATGGGTGCAGGTGATATTCAAAA 600

Qy   1266 ATTACAAAAATGCATATTTAGATAAAATGCATGAAAAAAATGCGTTTTAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATTACAAAAATGCATATTTAGATAAAATGCATGAAAAAAATGCGTTTTAA 649

RESULT 14
BD178031
LOCUS       BD178031        660 bp    DNA     linear   PAT 16-APR-2003
DEFINITION  MurC.
ACCESSION   BD178031
VERSION     BD178031.1  GI:30015295
KEYWORDS    JP 2002300888-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1 (bases 1 to 660)
  AUTHORS   Wallis,N.G. and Burnham,M.K.R.
  TITLE     MurC
```

```
Qy      1266 ATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTAA 1314
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTAA 649

RESULT 15
AE017039/c
LOCUS       AE017039              291804 bp    DNA     linear   BCT 30-APR-2003
DEFINITION  Bacillus anthracis str. Ames section 16 of 18 of the complete
            genome.
ACCESSION   AE017039 AE016879
VERSION     AE017039.1  GI:30259317
KEYWORDS    .
SOURCE      Bacillus anthracis str. Ames
  ORGANISM  Bacillus anthracis str. Ames
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 291804)
  AUTHORS   Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I.,
            Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E.,
            Okstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M.,
            Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S.,
            Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H.,
            Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J.,
            Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A.,
            Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S.,
            Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and
            Fraser,C.
  TITLE     The genome sequence of Bacillus anthracis Ames and comparison to
            closely related bacteria
  JOURNAL   Nature 423 (6935), 81-86 (2003)
   MEDLINE   22608414
   PUBMED    12721629
REFERENCE   2  (bases 1 to 291804)
  AUTHORS   Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I.,
            Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E.,
            Okstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M.,
            Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S.,
            Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H.,
            Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J.,
            Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A.,
            Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S.,
            Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and
            Fraser,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAR-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..291804
                     /organism="Bacillus anthracis str. Ames"
                     /mol_type="genomic DNA"
                     /strain="Ames"
                     /db_xref="taxon:198094"
     gene            complement(119..475)
                     /gene="rplT"
                     /locus_tag="BA4817"
     CDS             complement(119..475)
                     /gene="rplT"
                     /note="identified by match to PFAM protein family HMM
                     PF00453"
                     /codon_start=1
                     /transl_table=11
                     /product="ribosomal protein L20"
                     /protein_id="AAP28506.1"
                     /db_xref="GI:30259318"
                     /translation="MPRVKGGTVTRQRRKKVIKLAKGYYGSKNTLFKVANQQVMKSLM
                     YAFRDRRQKKRDFRKLWITRINAAARMNGLSYSRLMHGLKNAGIEVNRKMLADLAVHD
                     EKAFAELATVAKNNIN"
     gene            complement(513..713)
                     /gene="rpmI"
                     /locus_tag="BA4818"
     CDS             complement(513..713)
                     /gene="rpmI"
                     /locus_tag="BA4818"
                     /note="similar to SP:P55874; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="ribosomal protein L35"
                     /protein_id="AAP28507.1"
                     /db_xref="GI:30259319"
                     /translation="MPKQKTHRGAAKRFKKTGSGKLKRSHAYTSHLFANKSTKAKRKL
                     RKAGVVSAGDFKRIRQMLDNLK"
     gene            complement(735..1295)
                     /gene="infC"
                     /locus_tag="BA4819"
     CDS             complement(735..1295)
                     /gene="infC"
                     /locus_tag="BA4819"
                     /note="similar to SP:P55872; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="translation initiation factor IF-3"
                     /protein_id="AAP28508.1"
                     /db_xref="GI:30259320"
                     /translation="MTTNYVLRKPWRWLTISKDMINEQIRAREVRLVGANGDQLGIK
                     SRNDALDLAANLNLDLVLVAPNAKPVCRIMDYGKFRFEQQKKEKEQKRNQKVISMKE
                     VRLSPTIDEHDPNTKLRNAIKFLEKGDKVKASIRFKGRAITHKEIGQRVLDRFSEACA
                     EVSTIESKPKMBGRSMFLVLAPKNDK"
     gene            complement(1632..3569)
                     /gene="thrS-2"
                     /locus_tag="BA4820"
     CDS             complement(1632..3569)
                     /gene="thrS-2"
                     /locus_tag="BA4820"
                     /note="similar to SP:P18255; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="threonyl-tRNA synthetase"
                     /protein_id="AAP28509.1"
                     /db_xref="GI:30259321"
                     /translation="MADVVKITFPDGAVKEFPKGVTTEIAASISPGLKKKAVAGKLN
                     DEMIDLVTPIEEDGAVSIITLDSBDGLYILRHSTAHLLAQALKRLYKDVKVELGIGPV
                     IENGFYYDIDMEEAITVEDFKKIEKEMQKIVNENLEIVRHEVPRAEAIRRFEEIGDEL
                     KLDLINDLPEDAVISIYEQGEFFDLCRGVHLPSTGKIVFKLLSVAGAYWRDGNNKLN
                     LQRIYGTAFVKKAELDEHLRMLEEAKERDHRKLGKELKLFTNSQKVQCGLPLWLPKGA
                     TIRRIIERYIVDKEASLGYDHVYTPVLGSRELYETSGHWNHYRDGMPFSMEMDNEELV
                     LRPMNCPHHMVYKNDIHSYRELPIRIAELQTWHYREMSGALSGLRPVRGMTLNDAHI
                     FVRPDQIKEELKRVNLTLEVYKDFGLENYSFRLSYRDPEDTKKYYADDEMWEKAQGM
                     LKEAMDEMGLDYYEAEGEAAFYGPKLDVQVRTALGKDETLSTVQLDFLLPERFELSYV
                     GEDGKQHRPVVIHRGVVSTMERFVAFLIEBYKGAFPTWLAPVQVQVIPVSPQVHLDYA
                     KKVQDELRRAGIRVELDTREEKIGYKIREAQMQKIPYMLVVGDNEVTENGVNVRKYGE
                     QKSETIALDAFVDMIKVEGKR"
     gene            complement(3876..4757)
                     /locus_tag="BA4821"
     CDS             complement(3876..4757)
                     /locus_tag="BA4821"
                     /note="similar to GB:J05243, GB:M18627, GB:M19725,
                     GB:M24773, GB:X86901, GB:U26396, SP:Q13813, PID:178426,
                     PID:179106, PID:1805280, PID:537331, PID:829111,
                     PID:836669, GB:J05243, GB:M18627, GB:M19725, GB:M24773,
                     GB:X86901, GB:U26396, SP:Q13813, PID:178426, PID:179106,
                     PID:1805280, PID:537331, PID:829111, and PID:836669;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAP28510.1"
                     /db_xref="GI:30259322"
                     /translation="MIEICFEKNDAMHVVRQLLKRAELLYKETSVILQDQKVVIHIP
                     VCESNYIEKILLPVMVYFIVNVKQNEWIYTILKEKFYEBEBECHQILHMAQEILKGR
                     RKGIARELTRHTFESYIKSSLNNWLCDPLSFSFSSSYVRFRLRTYREMVAKLAEVAIDE"
```

```
Db  103051 CTTCATGACATGAAGCATACTGTTCAAGGGTCTGATTATGAAAAGCGTTCTTCTTCACACAA 102992
                          YKLKEQEYQMFIETLRQQVRSRKSRLSCVHLIFDESFIFYDDKGRRLKQEKLVRYMDED
                          LLKQKDVVIDTKVTAPLLSISPKKIYLYTKEQDHNMIITLRNVPQERVQLHGLHRFER
                          NVRNLKNKGNALDFLSF"

gene    complement(5008..5976)
        /gene="dnaI"
Qy   121 GTTGCTCTTGAAATAAGGGGATAAATATTACCATTGATGCTAATAACATAAAAGAA 180
         ||||||||||  |||||| || | |||| |||||  ||||||||||  |||||||
Db  102991 ACAGCCTTGGAAAAGCCTAATATCTCCTTCCTTTGATAAAAGTAATGTAAAAGAA 102932

CDS     complement(5008..5976)
        /gene="dnaI"
Qy   181 GATATGGTAGTTATACAAGGTAATGCATCGGCGAGTAGCCATGAGAAGAAATAGTACGTGCA 240
         ||||||| ||||||||||||||  || |||||||||||  | |||||||||||||
Db  102931 GGACAAGTGATTATTGCAGGAAATGCATTCCTGATACGCATGAAGAAATCGTAGCAGCA 102872

         /note="similar to GB:X57110, GB:X69207, SP:P22681,
         PID:29731, GB:X57110, GB:X69207, SP:P22681, and
         PID:29731; identified by sequence similarity; putative"
         /codon_start=1
         /transl_table=11
Qy   241 CATCAATTGAAATTAGTATGTGTAAGTATATATGATTTTTAGGACAGAGATTATTGATCAA 300
         |||||||| || || |  ||||  ||| ||  |  ||||||
Db  102871 AAAGAATTAAAACATCCCAGTACATCGTTACCACTACACTTCTTAGGTGATCTTATGAACCAA 102812

         /product="primosomal protein DnaI"
         /protein_id="AAP28511.1"
         /db_xref="GI:30259323"
Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATTCA 360
         |||||||||||| |  || ||||||||  ||  || ||| ||| ||| |||||
Db  102811 TACACAAGTGTTGCTGTAACTGGTGCGCATGGTAAAAAACATCAACAACTGGTTTGTTAGCC 102752

         /translation="MEHIQNSFAKLMENENFKNRYEVLKABVMAHPRVKEFIDEHRGE
         VTTSMIERSLVKLYEYIGQSVGCADCPDLGSCKNMLQGYTEPKLVIQKMDIQYDRCV
Qy   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGATGGCACAGGTATGGGA 420
         |||| || || ||  || |||  |   || ||||  | || || ||  ||||| || |
Db  102751 CATGTAATGCAAGGTGCACACCCTACCTATTGGAGATGAACAGGGCATGGG 102692

         RKVAYDERKKYEKLVQSVYMPTDILQATMENLDPSDLDARIDAIGAANEFLSAYEPGK
         KVQGLYLYGKFGVGKTYLLGAINELARKKISSMLVYPFPFLREIKSSIQDNSIGEKI
         DAVKRVQVLMLDDIGAEAMSSFVRDDVLGAILQFRMLENLPTFFTSNFDFKQLEHHLT
         YTQRGEAEEMKAARIMERIKYLAKPIPICGKNRRHK"
Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATGACGTCACTTTTTAAGT 480
         |||||| |||| | ||||||  ||| |||||| || |||||| |||| |||||||| |
Db  102691 GTAGAAAATAGTAAGTATTTGTATTTGTGAGTATCGTCGTCATTTTCTTGTCT 102632

gene    complement(6010..7416)
        /gene="dnaB"
Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTAATTTTAAA 540
         |||||| |||||||| ||||||||||| |||||||||| |||||| ||||||||||
Db  102631 TACAATCCAGACTATGCAATTATGACAAACATTGATTTTGATCATCCAGATTATTTCACA 102572

CDS     complement(6010..7416)
        /gene="dnaB"
        /codon_start=1
        /transl_table=11
Qy   541 GATATTAATGATGTTTTGATGATGAACATCACGTAAAATTAAAATGGCACATAATGTTAAAAAAGTATT 600
         |||| ||||| |||| | |||||||||||| |||  | ||||||||||  || ||||
Db  102571 GATATCAATGATGTTTCAGTGCATTCCAAGAGATGGCATTGCAAGTGAAAAAAGCATT 102512

         /product="DNA replication protein DnaB"
         /protein_id="AAP28512.1"
         /db_xref="GI:30259324"
Qy   601 ATTGCTTGGGTGATGATGAACATCACGTAAAATCACGTAAATGTTCCAATTTATTAT 660
         ||||||||| ||||||||| |||| ||| | |||||  |||| ||||||  ||||
Db  102511 ATTGCATGTGGAGATGATGAAGAACTTCAAGCGAAAAATTCAAGGTACCTGTTATTTTC 102452

         /translation="MEKSWMELLPIDRYKVSAKGLLHNYDRKVLTMLYQPLIGSRAF
         SLYMTIWGELEQDRVFGKENTHHSLMVTMQMQLPEVYEBRVKLEAIGLLKVYIKKEKD
         IRMFIYELQPPLSPKQFDDIVLSIFLYNRLSRTKYNQVKHYFLEEEFDPASYENVTR
         SFNDVFGSFNPGLEHAQEDLRIPKTTAMPSNEKGDAPKVWNDFFDFSLFVDGLSALV
Qy   661 TATTGGATTAAAGATTCGGATGACATTTATGCTCAAAAATATTCAAATTACGGATAAAGGT 720
         |||||| |||| || || ||| ||| |||| |||| ||| ||||| | ||| ||| ||
Db  102451 TATGGATTTGGAAGATAATGATTTCCAAGCACGTAACAATTCAAAAGAGAACTCACGGT 102392

         PKKAITDQVRECVITLAYVYDVDVLSMONIVLGAVTEMQTIDMERLRKGARDWYQFEN
         GQALPVLSERVQPHAARMKEKEPSTQEEMLIKQLEEISPKQLLKEISGAEATKADL
         QIVEDVMNQKITPGVVMLIYVMLRSDMKLAKTYVEKIAGHWARKVGTVARAMAL
         AKEENRQYQEWAETKKKGRTSKKTVRKEMVPDWLKERPKEQEKETVKKDASABKGAST
         LEDERKRLEEVLKKYKRD"
Qy   721 ACTGCTTTTGATGTGTATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||| |||||| ||| || |||||| | |||  ||| ||||  ||||||||
Db  102391 ACTATTTTCGATGTATTCGTTCGTAATACGTACTATGACACGTTCAAAATTACAGGCTAC 102332

gene    complement(7544..7969)
        /gene="speD-1"
        /locus_tag="BA4824"
Qy   781 GGTGACCATACAGGTTTAAATGCATTAAATGCGTAAATGCCACTTGCCATTAGAGAAGCTA 840
         ||| ||  |||  |||  |||| ||| ||||| |||| |||||| ||||||||  |||
Db  102331 GGCAACCACAGCGTATTAAATGCATTAGCAGTAGCATGCCACTTGCCATTATGAAAACGTT 102272

CDS     complement(7544..7969)
        /gene="speD-1"
        /note="identified by match to PFAM protein family HMM
        PF03477"
        /codon_start=1
        /transl_table=11
Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTCGTGTTAAACGTCGTTTCAAT 900
         ||||||||| | ||| |||| || ||| |||| ||  |||  ||  |||  |||| |
Db  102271 GATGTAGAAGCAGTTAAAGGTAACAACTTTGAGAGGCGTAAAACGTCGCTTTAAT 102212

         /product="ATP cone domain protein"
         /protein_id="AAP28513.1"
         /db_xref="GI:30259325"
Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGAGAATATCGAAAGAAATT 960
         |||| || |||||||||||||| |||| |||| |||| |||||||||
Db  102211 GAAAAGCCAATGGAGAGCAAGTTATTATTGTGACTACGCAGAACCACCAGCAGAAATT 102152

         /translation="MLDSRPVDEGRSIRRRECESCLSRFTTFERVEESPLIVVKKEG
         TREEFNEKILRGLLIKACEKRPVSLRQLEEVTQSVERERLRNLGISEVKSDMIGBIVME
         ELRDIDDVAYRFASVYRQFKDLNVFIEELKDILQKERE"
gene    complement(8291..8683)
        /gene="speD-1"
        /locus_tag="BA4825"
Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         |||| || ||||||| |||||||  ||||||| ||||||||| || |||| ||||| |
Db  102151 AATGCAACGATTGAAGCAGCTGTCGTCAAAAACATCCAGACGCGTGAAATTGTCGCTGTATTC 102092

CDS     complement(8291..8683)
        /gene="speD-1"
Qy   1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAAATTGCAGAAAGTTTAAGT 1080
          |||| ||| |||| | |||||||| || |||  |||| |||||| ||||||||  ||
Db  102091 CAGCCGGACCACATTCCACGTACAGAAAAAGTTCTTAGATGAGTTCGCTGAAAAGCCTAAGC 102032

Qy    1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACCAAATC 60
         |||| ||  |||| |||| ||| |||||||  ||||||||||||| ||||| |||  |||
Db  103111 ATGACAGTTTACCATTTTGTAGGAATTAAAGGAACACAGGAATTCATTAGCCGCAAATT 103052

Qy   61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTCGGATAACGTGATATTTACAGAA 120
         ||||| ||||||  ||||| ||||||||  ||| || |  ||||  ||||| |||||  ||
Db  101971 TTAACAATCGAAGATCTGCAAAAGCGTATTGACGGTGCAGAACTAATTACAGATACAACA 101912
```

```
Qy     1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATTATGGGTGCAGGTGATATT 1260
            |  ||||||||| |  ||  || |||||  | | ||||| ||| |||| ||||  |  |||||||
Db   101911 ACGGATGTATTAAAGAAACATAAAAACGGCGTTCTCATTTTCATGGGCGCAGGAGACATC 101852

Qy     1261 CAAAAATTACAAAATGCATATTTAGATAAATT 1292
            ||||||||  |  ||  | ||| |  || |  |
Db   101851 CAAAAATTCGAAGCAGCAGCTTACGTAAAGAAGT 101820
```

Search completed: June 6, 2004, 01:40:50
Job time : 5358 secs

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   June  6, 2004, 02:44:45 ; Search time 635 Seconds
                              (without alignments)
                              8817.520 Million cell updates/sec

Title:         US-10-712-713-1
Perfect score: 1318
Sequence:      1 atgacacactatcatttgt........gaaaaatgcgtttaagctt 1318

Scoring table: OLIGO_NUC
          Gapop 60.0 , Gapext 60.0

Searched:     3373863 seqs, 2124099041 residues

Word size :   50

Total number of hits satisfying chosen parameters:    35

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
         1:  geneseqn1980s:*
         2:  geneseqn1990s:*
         3:  geneseqn2000s:*
         4:  geneseqn2001as:*
         5:  geneseqn2001bs:*
         6:  geneseqn2002s:*
         7:  geneseqn2003as:*
         8:  geneseqn2003bs:*
         9:  geneseqn2003cs:*
        10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Match	Query Length	DB	ID	Description
1		1318	100.0	1318	4	AAS00189	Aas00189 S. aureus
2		1314	99.7	1335	9	ADD15462	Add15462 Staphyloc
3		1311	99.5	1332	9	ACF72770	Acf72770 Staphyloc
4		1263	95.8	1335	9	ADD15464	Add15464 Staphyloc
5		1212	92.0	1314	7	ACA19726	Aca19726 Prokaryot
6		957	72.6	1335	4	AAS54558	Aas54558 Staphyloc
7		954	72.4	1311	4	AAS51660	Aas51660 Staphyloc
8		842	63.9	1351	2	AAV99650	Aav99650 UDP-N-ace
9		740	56.1	2424	2	AAV74703	Aav74703 Staphyloc
10	c	330	25.0	660	2	AAV80065	Aav80065 Partial n
11	c	275	20.9	619	2	AAV53479	Aav53479 DNA encod
12	c	257	19.5	257	4	AAS50351	Aas50351 Staphyloc
13	c	257	19.5	257	4	ACA17633	Aca17633 Prokaryot
14	c	225	17.1	225	4	AAS49022	Aas49022 Staphyloc
15	c	225	17.1	225	4	AAS49054	Aas49054 Staphyloc
16	c	225	17.1	225	4	AAS49060	Aas49060 Staphyloc
17	c	225	17.1	225	4	AAS48994	Aas48994 Staphyloc
18	c	225	17.1	225	7	ACA16308	Aca16308 Prokaryot
19	c	225	17.1	225	7	ACA16221	Aca16221 Prokaryot
20	c	225	17.1	225	7	ACA16304	Aca16304 Prokaryot
21	c	225	17.1	225	7	ACA16261	Aca16261 Prokaryot
22	c	206	15.6	206	4	AAS48839	Aas48839 Staphyloc
23	c	206	15.6	206	7	ACA16071	Aca16071 Prokaryot
24	c	168	12.7	227	7	ACA16230	Aca16230 Prokaryot
25	c	158	12.0	158	4	AAS51038	Aas51038 Staphyloc
26	c	158	12.0	158	7	AAS50998	Aas50998 Staphyloc
27	c	158	12.0	158	7	ACA18282	Aca18282 Prokaryot
28	c	158	10.8	158	4	AAS50478	Aas50478 Staphyloc
29	c	143	10.8	143	7	ACA17740	Aca17740 Prokaryot
30	c	143	10.8	143	3	AAA26910	Aaa26910 Essential
31	c	90	6.8	888	4	AAF91592	Aaf91592 Staphyloc
32	c	90	6.8	888	4	AAS08061	Aas08061 Staphyloc
33	c	90	6.8	888	9	ADD67439	Add67439 Antibacte
34	c	90	6.8	888	9	ADE73663	Ade73663 Mutant ba
35	c	90	6.8	888			

ALIGNMENTS

```
RESULT 1
AAS00189
ID   AAS00189 standard; DNA; 1318 BP.
XX
AC   AAS00189;
XX
DT   04-JUL-2001  (first entry)
XX
DE   S. aureus DNA encoding UDP-N-acetylmuramate:L-alanine ligase, MurC.
XX
KW   UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
KW   antibody; wound infection; cellulitis; burn infection; eyelid infection;
KW   food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW   skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW   Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW   ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1314
FT                   /*tag= a
FT                   /product= "MurC"
XX
PN   WO200116292-A2.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US023773.
XX
PR   01-SEP-1999; 99US-0151933P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Choi GH;
XX
DR   WPI; 2001-183259/18.
DR   P-PSDB; AAU00828.
XX
PT   New isolated nucleic acid for use in diagnosing Staphylococcus infections
PT   and in vaccines for eliciting immune responses to the infections.
XX
PS   Claim 1; Page 14; 225pp; English.
XX
CC   The sequence encodes S. aureus MurC (UDP-N-acetylmuramate:L-alanine
CC   ligase). The polynucleotides of the invention are used to detect
CC   Staphylococcus nucleic acids in a biological sample from an animal for
CC   diagnosing Staphylococcus infections. The polypeptides of the invention
CC   are used to detect anti-Staphylococcus antibodies in a biological sample
CC   from an animal to diagnose Staphylococcus infections. The polypeptides
CC   are also used in vaccines to elicit protective antibodies in an animal to
CC   a member of the Staphylococcus genus and for preventing or attenuating an
CC   infection caused by a member of the Staphylococcus genus e.g wound
CC   infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC   joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC   scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
```

```
CC   disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC   polynucleotides may also be used in vaccines and for preventing or
CC   attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC   may be used to purify, detect and target the polypeptides in in vitro and
CC   in vivo diagnostic and therapeutic methods
XX
SQ   Sequence 1318 BP; 464 A; 162 C; 251 G; 441 T; 0 U; 0 Other;

Query Match        100.0%;  Score 1318;  DB 4;  Length 1318;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1318;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGACACACTACCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGACACACTACCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC  60

Qy     61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA  120

Qy    121 GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  180

Qy    181 GATATGGTAGTTTATACAAGTAATGCATTCGGCGAGTAGCCATGAGAGAAATAGTAGTGCA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATATGGTAGTTTATACAAGTAATGCATTCGGCGAGTAGCCATGAGAGAAATAGTAGTGCA  240

Qy    241 CATCAATTGAAATTAGATGTGTGAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CATCAATTGAAATTAGATGTGTGAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA  300

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACACACAGGTTTATTATCA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACACACAGGTTTATTATCA  360

Qy    361 CATGTTATGAAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATGTTATGAAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA  420

Qy    421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATATAGAGACGTCACTTTTTAAGT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTGCCTGAAAGTGATTATTCGCTTTGAGGCATGTGAATATAGAGACGTCACTTTTTAAGT  480

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  600

Qy    601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660

Qy    661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT  720

Qy    721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  780

Qy    781 GGTGACCATACAGTTTAAATGCTAATTGCGATTAGTATTGCAGAAGCTA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGTGACCATACAGTTTAAATGCTAATTGCGATTAGTATTGCAGAAGCTA  840

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGCGTGTTGTAAACGTCGTTCAAT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGCGTGTTGTAAACGTCGTTCAAT  900

Qy    901 GAAACTACAATTGCAAATCAGTCGCACACCATCCAAGAGAAATT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAACTACAATTGCAAATCAGTCGCACACCATCCAAGAGAAATT  960

Qy    961 AGTTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGAAGTGTTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGTTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGAAGTGTTGCAGTATTT 1020

Qy   1021 CAACCACCACACTTCTCTAGAACACACAGGCATTTTAAATGAATTGCAGAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAACCACCACACTTCTCTAGAACACACAGGCATTTTAAATGAATTGCAGAAGTTTAAGT 1080

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCA 1140

Qy   1141 TTAACGATACAAGATTAATTGATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTAACGATACAAGATTAATTGATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200

Qy   1201 ATTAATGTATTAGAACAATTGATAAGTGCTGTTATTTATTTATGGGTGCAGGTGATATT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ATTAATGTATTAGAACAATTGATAAGTGCTGTTATTTATTTATGGGTGCAGGTGATATT 1260

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATAATTAGATAAATAGGCATGAAAAATGCGTTTTAAGCTT 1318
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CAAAAATTACAAAATGCATATTTAGATAAATAATTAGATAAATAGGCATGAAAAATGCGTTTTAAGCTT 1318

RESULT 2
ADD15462
ID   ADD15462 standard; DNA; 1335 BP.
XX
AC   ADD15462;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Staphylococcus aureus murC DNA (SeqID 27).
XX
KW   microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW   immunological response; vaccination; surface disinfectant;
KW   personal hygiene application; food preservative; ds; gene; murC;
KW   UDP-N-acetylmuramate-alanine ligase.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1335
FT                   /*tag= a
FT                   /product= "MurC protein"
XX
PN   WO2003025007-A2.
XX
PD   27-MAR-2003.
XX
PF   20-SEP-2002; 2002WO-CA001428.
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.
PR   18-DEC-2001; 2001US-0341732P.
PR   18-DEC-2001; 2001US-0341776P.
PR   19-DEC-2001; 2001US-0341949P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A,  Dharamsi A,  Vedadi M,  Awrey D,  Beattie B;
PI   Domagala M,  Houston S,  Kanagarajah D,  Nethery K,  Ng I,  Mansoury K;
PI   Mcdonald M,  Pinder B,  Viola C,  Wrezel O;
XX
```

DR   WPI; 2003-468119/44.
DR   P-PSDB; ADD15463.
XX
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 27; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polynucleotide
CC   sequence is DNA predicted from the genomic sequence of S. aureus UDP-N-
CC   acetylmuramate-alanine ligase (murC) of the invention.
XX
SQ   Sequence 1335 BP; 472 A; 161 C; 255 G; 447 T; 0 U; 0 Other;

Query Match   99.7%; Score 1314; DB 9; Length 1335;
  Best Local Similarity   100.0%; Pred. No. 0;
  Matches 1314; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ATGACACACTATCATTTGTCGGAATTAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATTTGTCGGAATTAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 81

Qy     61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATCGGATATTGAGAACTACGTATTACAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATCGGATATTGAGAACTACGTATTACAGAA 141

Qy    121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAGAA 201

Qy    181 GATATGGTAGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GATATGGTAGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCA 261

Qy    241 CATCAATTGAAATTAGAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGAGATGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 321

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 381

Qy    361 CATGGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 501

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTATGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 561

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 621

Qy    601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 681

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 741

Qy    721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATGCGATTAGTTATTAGAGAAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATGCGATTAGTTATTAGAGAAGCTA 861

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGGTGGTGTTAAACGTCGTTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGGTGGTGTTAAACGTCGTTTCAAT 921

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAATT 981

Qy    961 AGTTGCTACAATTGAAACAGCACGCAAAGAAAAATCCACCATAAAGAAGTTGTTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTTGCTACAATTGAAACAGCACGCAAAGAAAAATCCACCATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACCACACTTCTCTCAGAACACACAGGCATTTTAAATGAATTTGCAGAAAGTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACCACACTTCTCTCAGAACACACAGGCATTTTAAATGAATTTGCAGAAAGTTAAGT 1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTCATTGAGAAAAATACTGCGGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTCATTGAGAAAAATACTGCGGCA 1161

Qy   1141 TTAACGATACAAGATTAAATTGAATAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTAAATTGAATAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1221

Qy   1201 ATTAATGTATTAGACAATTGATAATGCTGTTATTTATTTATGGGTGCAGGTGATATT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGTATTAGACAATTGATAATGCTGTTATTTATTTATGGGTGCAGGTGATATT 1281

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1335
```

RESULT 3
ACF72770

```
Db    622 ATTGCTTGGGGTGATGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT  681
                                                                          
Qy    661 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT     720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT     741

Qy    721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT     780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT    801

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA      840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA      861

Qy    841 GATGTTACAAATATTAAAGAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT      900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAATATTAAAGAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT      921

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTAATGCACACCATCCAAGAGAAATT   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTAATGCACACCATCCAAGAGAATT   981

Qy    961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTGTTGTGCAGTATT   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTGTTGTGCAGTATT   1041

Qy   1021 CAACCACACCTTTCTCTAGAACACCAGGCATTTTTAAATGAATTTGCAGAAGTTTAAGT    1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACCTTTCTCTAGAACACCAGGCATTTTTAAATGAATTTGCAGAAGTTTAAGT    1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTCGGATCAATTAGAGAAAATACTCGGCGCA  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTCGGATCAATTAGAGAAAATACTCGGCGCA  1161

Qy   1141 TTAACGATACAAGATTAATTGATAAAATGAAGGTGCATCGTAATTAATGAAGATTCT     1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTAATTGATAAAATGAAGGTGCATCGTAATTAATGAAGATTCT     1221

Qy   1201 ATTAATGTATTAGAACAATTGATAATGCTGTTTATTTTATTATCGGTGCAGGTGATATT    1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGTATTAGAACAATTTGATAATGCTGTTTATTTTATTATCGGTGCAGGTGATATT   1281

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGCCATGAAAAAATGCGTTT           1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAAATGCATATTTAGATAAATTAGCCATGAAAAAATGCGTTT           1332
```

```
Query Match      99.5%;  Score 1311;  DB 7;  Length 1332;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1311;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGTAGCACAAATC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGTAGCACAAATC   81

Qy     61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTCGATATTTACAGAA          120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTCGATATTTACAGAA          141

Qy    121 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAA   201

Qy    181 GATATGGTAGTTATACAAGTAATGCATCGCGGAGTAGCCATGAAGAAAATAGTAGTACGTGCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GATATGGTAGTTATACAAGTAATGCATCGCGGAGTAGCCATGAAGAAAATAGTAGTACGTGCA 261

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGAGTAATTGATCAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGAGTAATTGATCAA  321

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAACTTCTACAACACAGGTTTATTATCA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAACTTCTACAACACAGGTTTATTATCA   381

Qy    361 CATGTTAATGAATGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGTTAATGAATGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA    441

Qy    421 TTGCTGAAAGTGATTATTCGCTTTTGAGGCATGGAATATAGACGTCACTTTTTAAGT    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCTGAAAGTGATTATTCGCTTTTGAGGCATGGAATATAGACGTCACTTTTTAAGT    501

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATGATTTCCATCATCCTGATTATTTTAAA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTACGCAATTATGACAAATATGATTTCCATCATCCTGATTATTTTAAA   561

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGCGACACATAATGTTAAAAAAGGTATT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGCGACACATAATGTTAAAAAAGGTATT  621

Qy    601 ATTGCTTGGGGTGATGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

PF   20-SEP-2002; 2002WO-CA001428
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.
PR   18-DEC-2001; 2001US-0341732P.
PR   18-DEC-2001; 2001US-0341776P.
PR   19-DEC-2001; 2001US-0341949P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A, Dharamsi A, Vedadi M, Alam MZ, Awrey D, Beattie B;
PI   Domagala M, Houston S, Kanagarajah D, Nethery K, Ng I, Mansoury K;
PI   Mcdonald M, Pinder B, Viola C, Wrezel O;
XX
DR   WPI; 2003-468119/44.
DR   P-PSDB; ADD15465.
XX
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 29; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polynucleotide
CC   sequence is the experimentally predicted DNA of S. aureus UDP-N-
CC   acetylmuramate-alanine ligase (murC) of the invention.
XX
SQ   Sequence 1335 BP; 472 A; 160 C; 255 G; 448 T; 0 U; 0 Other;

Query Match   95.8%; Score 1263; DB 9; Length 1335;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1313; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ATGACACACTATCATTTTGTCGGAATTAAAGTTCTGGCATGAGTTCATTAGCACAAATC   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATTTTGTCGGAATTAAAGTTCTGGCATGAGTTCATTAGCACAAATC   81

Qy   61 ATGACATGATTAGGACATGAGAGTTCAAGGATCCGGAGATATTGGAGAACTACGTATTTACAGAA   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGACATGATTAGGACATGAGAGTTCAAGGATCCGGAGATATTGGAGAACTACGTATTTACAGAA   141

Qy   121 GTTGCTCTTAGAAATAAGGATAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GTTGCTCTTAGAAATAAGGATAAAATATTACCATTTGATGCTAATAACATAAAAGAA   201

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAATAGTACGTGCA   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAATAGTACGTGCA   261

Qy   241 CATCAATTGAAATTAGATGTTGTAAGTTATAAATGATTTTTAGGACAGATTATTGATCAA   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CATCAATTGAAATTAGATGTTGTAAGTTATAAATGATTTTTAGGACAGATTATTGATCAA   321

Qy   301 TATACTTCAGTAGCTGTAACTGGTGGTGCACATGGTGTAAAACTTCTACAACAGGTTTATTATCA   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 TATACTTCAGTAGCTGTAACTGGTGGTGCACATGGTGTAAAACTTCTACAACAGGTTTATTATCA   381

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   441

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT   501

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA   561

Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT   621

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGAGATGTTCCAATTTATTAT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ATTGCTTGGGGTGATGATGAACATTACGTAAAATTGAAGCAGAGATGTTCCAATTTATTAT   681

Qy   661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT   741

Qy   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT   801

Qy   781 GGTGACCATACAGTTTAAATGCATTAAGTCGTAATTGCGATTAGTATTAGAGAAGCTA   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GGTGACCATACAGTTTAAATGCATTAAGTCGTAATTGCGATTAGTATTAGAGAAGCTA   861

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTTAAACGTCGTTTCAAT   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTTAAACGTCGTTTCAAT   921

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTGAGATGATGCACACCATCCAAGAGAAATT   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACAATTGCAAATCAAGTTATTGTGAGATGATGCACACCATCCAAGAGAAATT   981

Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATCCACATAAAGAAGTTGTTGCAGTATTT   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAATTGAAACAGCACGAAAGAAATCCACATAAAGAAGTTGTTGCAGTATTT   1041

Qy 1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1042 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT   1101

Qy 1081 AAAGCAGATGCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCA   1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1102 AAAGCAGATGCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCA   1161

Qy 1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1162 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1221

Qy 1201 ATTAATGTATTAGAACAATTGATAATGCGTTAATTTTATTATGGGTGCAGTGATATT   1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1222 ATTAATGTATTAGAACAATTGATAATGCTGTATTTTATTATGGGTGCAGTGATATT   1281

Qy 1261 CAAAAATTACAAAAATGCATATTAGATAAATTAGGCATGAAAATGCGTTTAA   1314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1282 CAAAAATTACAAAAATGCATATTAGATAAATTAGGCATGAAAATGCGTTTAA   1335

RESULT 5
ACA19726
ID   ACA19726 standard; DNA; 1314 BP.

```
XX
AC   ACA19726;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #1383.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus aureus.
XX
PN   WO2002771183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU15856.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 7596; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation or to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1314 BP; 464 A; 161 C; 250 G; 439 T; 0 U; 0 Other;
```

```
Best Local Similarity  99.8%;  Pred. No. 0;
Matches 1312;  Conservative  0;  Mismatches  2;  Indels  2;  Gaps  0;

Qy     1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60

Qy    61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAAACTACGTATTTACAGAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAAACTACGTATTTACAGAA 120

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACACATAAAAGAA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACACATAAAAGAA 180

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCCGCAGTAGCCATGAAGAAATAGTACGTGCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATATGGTAGTTATACAAGGTAATGCATTCCGCAGTAGCCATGAAGAAATAGTACGTGCA 240

Qy   241 CATCAATTGAAATTAGATGTTGTAAGTTATATAATGATTTTTTAGGACAGATTATTGATCAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCAATTGAAATTAGATGTTGTAAGTTATATAATGATTTTTTAGGACAGATTATTGATCAA 300

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATAGACGTCACTTTTTAAGT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATAGACGTCACTTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600

Qy   601 ATTGCTTGGGGTGATGATGAACATCTCACGTAAAAATTGAAGCAGATGTTCCAATTTATTAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGCTTGGGGTGATGATGAACATCTCACGTAAAAATTGAAGCAGATGTTCCAATTTATTAT 660

Qy   661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTTGATGTATGTGGATGGTAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTGCTTTTGATGTATGTGGATGGTAGTTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy   781 GGTGACCATACAGTTTAAATGCTGTAAATGCCATTAGTATTGCGATTAGAAGCTA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGTGACCATACAGTTTAAATGCTGTAAATGCCATTAGTATTGCGATTAGAAGCTA 840

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900

Qy   901 GAAACTACAAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAACTACAAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Qy   961 AGTGCTACACACTTCTCTAGAACACAGCACCATTTTAAATGAATTGCAGAAAGTTGTTGCAGTATTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGTGCTACACACTTCTCTAGAACACAGCACCATTTTAAATGAATTGCAGAAAGTTGTTGCAGTATTT 1020

Qy  1021 CAACCACACACTTCCTCTAGAACACAGCACCATTTTAAATGAATTGCAGAAAGTTAAGT 1080
```

```
XX
SQ   Sequence 1314 BP; 464 A; 161 C; 250 G; 439 T; 0 U; 0 Other;

Query Match   92.0%;   Score 1212;   DB 7;   Length 1314;
```

```
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1335 BP; 472 A; 164 C; 255 G; 444 T; 0 U; 0 Other;

Query Match             72.6%;  Score 957;  DB 4;  Length 1335;
Best Local Similarity   99.5%;  Pred. No.0;
Matches 1307;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        1 ATGACACACTATCATTTGTCCGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22 ATGACACACTATCATTTGTCCGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC   81

Qy       61 ATGCATGATTTAGGACATGAAGTTCAAGGATGCGGATATTGAGAACTACGTATTTACAGAA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       82 ATGCATGATTTAGGACATGAAGTTCAAGGATGCGGATATTGAGAACATACGTATTTACAGAA  141

Qy      121 GTTGCTCTTAGAGAAATAAGGGGATAAAATATTACCATTTGATGCTAATAACATAAAGAA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      142 GTTGCTCTTAGAGAAATAAGGGGATAAAATATTACCATTTGGTGCTAATAACATAAAAGAA  201

Qy      181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAAGAAATAGTACGTGCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      202 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAAGAAATAGTACGTGCA  261

Qy      241 CATCAATTGAAATTAGATGTGTGAAGTTATATGATGATTTTTAGGACGAGATTATTGATCAA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      262 CATCAATTGAAATTAGATGTGTGAAGTTATATGATGATTTTTAGGACACAGATTATTGATCAA  321

Qy      301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTCACAACAGGTTTATTATCA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTCACAACAGGTTTATTATCA  381

Qy      361 CATGTTTATGAATGGTGATAAAAAGACTTCATTTTTATTGGTGATGGCACAGGTATGGGA  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      382 CATGTTTATGAATGGTGATAAAAAGACTTCATTTTTATTGGTGATGGCACAGGTATGGGA  441

Qy      421 TTGCCTGAAAGTGATTGATTATTTCGCTTTTGAGGCATGTGAATATAGAGCGCACTTTTAAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      442 TTGCCTGAAAGTGATTGATTATTTCGCTTTTGAGGCATGTGAATATAGAGCGCACTTTTTAAGT  501

Qy      481 TATAAACCTGAGCTGATTACGGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTAAA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      502 TATAAACCTGAGCTGATTACGGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTCAA  561

Qy      541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACAATAATGTTAAAAAAGGTATT  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACAATAATGTTAAAAAAGGTATT  621

Qy      601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATGAAGCAGATGTTCCAATTTATTAT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      622 ATTGCTTGGGGTGATGATGAACATCACGTAAAATGAAGCAGATGTTCCAATTTATTAC  681

Qy      661 TATGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATATTACGGATAAAGGT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      682 TATGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATATTACGGATAAAGGT  741

Qy      721 ACTGCTTTGATGTGATGTGGATGGTGAGTTTTATGATCACTTCCGTCTCCACAATAT  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      742 ACTGCTTTGATGTGATGTGGATGGTGAGTTTTATGATCACTTCCGTCTCCACAATAT  801

Qy      781 GGTGACCATACAGTTTAAATGCGTAATTGCGATTAGTTATTAGAAGAAGCTA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      802 GGTGACCATACAGTTTAAATGCGTAATTGCGATTAGTTATTAGAAGAAGCTA  861

Qy      841 GATGTTACAAATATTAAAGAACATTAGAAACGTTGGTCGTGTTTTCAAT  900
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      862 GATGTTACAAATATTAAAGAACATTAGAAACGTTGGTCGTGTTTTCAAT  921
```

```
Db     1021 CAACCACACACTTCTCTAGAACACAAGCATTTTTAAATGAATTGCAGAAAGTTTAAGT  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1081 AAAGCAGATGCGTGTATTCTTATGTGAATTTTGGATCAATTAGAGAAAATACTGGCGCA  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 AAAGCAGATGCGTGTATTCTTATGTGAATTTTGGATCAATTAGAGAAAATACTGGCGCA  1140

Qy     1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200

Qy     1201 ATTAATGTTATTAGAACAATTGATAATGCTGTTATTTTATTTATTTATGGGTGCAGGTGATATT  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 ATTAATGTTATTAGAACAATTGATAATGCTGTTATTTTATTTATTTATGGGTGCAGGTGATATT  1260

Qy     1261 CAAAAATTACAAAAATGCATATTTAGATAAATTAGGCATGAAAAAATGCCGTTTTAA  1314
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CAAAAATTACAAAAATGCATATTTAGATAAATTAGGCATGAAAAAATGCCGTTTTAA  1314
```

```
RESULT 6
AAS54558
ID   AAS54558 standard; DNA; 1335 BP.
XX
AC   AAS54558;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #870.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zysking JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36699.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 8195; 51pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
```

CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1311 BP; 462 A; 164 C; 250 G; 435 T; 0 U; 0 Other;

Query Match   72.4%;   Score 954;   DB 4;   Length 1311;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 1304;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

Qy   1 ATGACACCACTACCATTTTGTCGGAATTAAAGGTTCTCGGCATGAGTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGACACCACTACCATTTTGTCGGAATTAAAGGTTCTCGGCATGAGTCATTAGCACAAATC 60

Qy   61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGTGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGGTGTGCTAATAACATAAAAGAA 180

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAAATAGTAGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAAATAGTACGTGCA 240

Qy   241 CATCAATTGAAATTAGATGTGTTAAGTATAATGATTTTTAGGACAGATTATTGATCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCAATTGAAATTAGATGTGTTAAGTATAATGATTTTTAGGACAGATTATTGATCAA 300

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGTTTATTATCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGTTTATTATCA 360

Qy   361 CATGTTAATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CATGTTAATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTCAAA 540

Qy   541 GATATTAATGATGTTTTGATGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTGATGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600

Qy   601 ACTTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCAATTTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAC 660

Qy   661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCCACAATAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACAATTGCAAATCAAGTTATGCGATGATTATGCACACCATCCAAGAGAAATT 981

Qy   961 AGTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAATTGAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACACACTTTCTCGAGACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCGAGACACAGCAATTTTTAAATGAATTGCAGAAAGTTTAAGT 1101

Qy   1081 AAAGCAGATGCTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATGCTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1161

Qy   1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1221

Qy   1201 ATTAATGTATTAGAAACAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGTATTAGAAACAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATATT 1281

Qy   1261 CAAAAAATTACAAAAATGCATATTTAGATAAATTAGCATGAAAAATGCGTTTTAA 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAAATTACAAAAATGCATATTTAGATAAATTAGCATGAAAAATGCGTTTTAA 1335

RESULT 7
AAS51660
ID   AAS51660 standard; DNA; 1311 BP.
XX
AC   AAS51660;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #77.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33801.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 4242; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,

```
PI   Wallis NG, Burnham MKR;
XX
DR   WPI; 1999-062655/06.
DR   P-PSDB; AAW87771.
XX
PT   New isolated MurC polypeptide from Staphylococcus aureus and related
PT   nucleic acid - useful in diagnosis, treatment and prevention of bacterial
PT   infections.
XX
PS   Claim 2; Page 3-4; 39pp; English.
XX
CC   The present sequence represents a MurC gene encoding a Staphylococcus
CC   aureus UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide). Host
CC   cell containing an expression system comprising the MurC gene can be used
CC   for the recombinant production of the polypeptide. Agonists or the MurC
CC   polypeptide are used to treat conditions requiring increased activity or
CC   expression of the polypeptide. Antagonists, inhibitory nucleic acid or
CC   competitive polypeptide are useful for inhibiting the polypeptide e.g.
CC   bacterial (especially S. aureus) infections. They are also useful against
CC   Helicobacter pylori infections and related cancers, ulcers and gastritis.
CC   The antibacterial agents are useful to treat in-dwelling devices for
CC   infection prevention or generally as wound treatments to prevent adhesion
CC   of bacteria to matrix proteins. The MurC polypeptide is also useful for
CC   diagnosing or prognosing a (susceptibility to) disease, for raising
CC   antibodies; to identify modulators or specific receptors; in rational
CC   drug design and as an immunogen for vaccines. The MurC gene sequences are
CC   useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
CC   for chromosomal mapping; to determine bacterial serotype; and for genetic
CC   immunisation
XX
SQ   Sequence 1351 BP; 471 A; 166 C; 260 G; 454 T; 0 U; 0 Other;

Query Match             63.9%;  Score 842;  DB 2;  Length 1351;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 1042;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy          1 ATGACACCATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGACACCATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 81

Qy         61 ATGATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ATGATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 141

Qy        121 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTGATGCTAATAACATAAAAGAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GTTGCTCCTAGAAATAAGGGGATAAAAAATATTACCATTGGTGCTAATAACATAAAAGAA 201

Qy        181 GATATGGTAGTTATACAAGGTAATGCATTCGGCAGTAGCCATGAAGAAATAGTACGTGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GATATGGTAGTTATACAAGGTAATGCATTCGGCAGTAGCCATGAAGAAATAGTACGTGCA 261

Qy        241 CATCAATTGAAATTAGAATTGTAAGTTATAATGATTATTGATCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        262 CATCAATTGAAATTAGAATTGTAAGTTATAATGATTATTGATCAA 321

Qy        301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 381

Qy        361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy        421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTTCACTTTTAAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTTCACTTTTAAGT 501

Qy        481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTCAAA 561

Qy        541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
```

```
Db        721 ACTGCTTTGATGTGTATGTCGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTAGTTATTAGAGAAGCTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTAGTTATTAGAGAAGCTA 840

Qy        841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTAAACGTCGTTTCAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTAAACGTCGTTTCAAT 900

Qy        901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Qy        961 AGTTGCTACATTGAAACAGCACGCGAAAGAAATATCCACATAAAGAAGTTGTGTTGCAGTATTT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGTTGCTACATTGAAACAGCACGCGAAAGAAATATCCACATAAAGAAGTTGTGTGCAGTATTT 1020

Qy       1021 CAACCACACACTTTCCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAACCACACACTTTCCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTTAAGT 1080

Qy       1081 AAAGCAGATCGTGATTCCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCGCA 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAAGCAGATCGTGATTCCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCGCA 1140

Qy       1141 TTAACGATACAAGATTAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTAACGATACAAGATTAAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200

Qy       1201 ATTAATGTATTAGAACAATTTGATAATGCTGTATTTATTTATTTATGGGTGCAGTGATATT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTAATGTATTAGAACAATTTGATAATGCTGTATTTATTTATTTATGGGTGCAGTGATATT 1260

Qy       1261 CAAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAAATGCGTTT 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CAAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAAATGCGTTT 1311
```

```
RESULT 8
AAV99650
ID   AAV99650 standard; DNA; 1351 BP.
XX
AC   AAV99650;
XX
DT   17-MAR-1999 (first entry)
XX
DE   UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide) encoding DNA.
XX
KW   MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW   bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW   immunogen; drug; genetic immunisation; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key            Location/Qualifiers
FT   CDS            22..1335
FT                  /tag= a
FT                  /gene= "MurC"
FT                  /product= "UDP-N-acetylmuramate:L-alanine ligase (MurC
FT                  polypeptide)"
XX
PN   EP889123-A2.
XX
PD   07-JAN-1999.
XX
PF   26-JUN-1998;  98EP-00305064.
XX
PR   03-JUL-1997;  97US-0052720P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
```

DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 1287-1288; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 2424 BP; 839 A; 302 C; 462 G; 759 T; 0 U; 62 Other;

Query Match   56.1%; Score 740; DB 2; Length 2424;
Best Local Similarity 100.0%; Pred. No. 0;
  Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   575 TGGCACATAATGTTAAAAAGGTATTATTGCTTGGGGTGATGATGAACATCTACGTAAAA 634
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGGCACATAATGTTAAAAAGGTATTATTGCTTGGGGTGATGATGAACATCTACGTAAAA 1260

Qy   635 TTGAAGCAGATGTTCCAATTATTATTATTATGGATTTAAAGATTCGGATGACATTTATGCTC 694
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTGAAGCAGATGTTCCAATTATTATTATTATGGATTTAAAGATTCGGATGACATTTATGCTC 1320

Qy   695 AAAATATTCAAATTACGGATAAAGGTACTCGCTTTTGATGTGTATGTGGATGGTGAGTTTT 754
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAAATATTCAAATTACGGATAAAGGTACTCGCTTTTGATGTGTATGTGGATGGTGAGTTTT 1380

Qy   755 ATGATCACTTCCTGTCTCCACAATATGGTACCATACAGTTTTAAATGCATTAGCTGTAA 814
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ATGATCACTTCCTGTCTCCACAATATGGTACCATACAGTTTTAAATGCATTAGCTGTAA 1440

Qy   815 TTGCGGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGT 874
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTGCGGATTAGTTATTTAGAGAAGCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGT 1500

Qy   875 TTGGTGGTGGTTAAACGTCGTTCAATGAAGAAGCAAATCAAGTTATTGTAGATG 934
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTGGTGGTGGTTAAACGTCGTTCAATGAAGAAGCAAATCAAGTTATTGTAGATG 1560

Qy   935 ATTATGCACACCATCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATC 994
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ATTATGCACACCATCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATC 1620

Qy   995 CACATAAAGAAGTTGTTGCAGTATTTCAACCACACTTTCTCAGAACACAGGCATTTT 1054
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CACATAAAGAAGTTGTTGCAGTATTTCAACCACACTTTCTCAGAACACAGGCATTTT 1680

Qy   1055 TAAATGAATTTGCAGAAAGTTTAGTAAAGTTTAGTAAAGTGTGTATTCTTATGTGAAATTTTTG 1114
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TAAATGAATTTGCAGAAAGTTTAGTAAAGTTTAGTAAAGTCGTGTATTCTTATGTGAAATTTTTG 1740

Qy   1115 GATCAATTAGAGAAAATACTGGGCATTAACGATCAAGATTAATTGATGAAATTGAAG 1174
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GATCAATTAGAGAAAATACTGGGCGCCATTAACGATCAAGATTAATTGATGAAATTGAAG 1800

Db   562 GATATTAATGATGATGTTTTTGATGATGAACATCATTCCAAGAAGAATGGGCACATAATGTTAAAAAAGGTATT 621

Qy   601 ATTGCTTGGGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTTAT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ATTGCTTGGGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTTAC 681

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 741

Qy   721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGCTTTGATGTGTATGTGGATGGTCAGTTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy   781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGGATTAGTTATTTAGAGAAGCTA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGGATTAGTTATTTAGAGAAGCTA 861

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGGTTAAACGTCGTTCAAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGGTTAAACGTCGTTCAAT 921

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 981

Qy   961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACACTTTCTCTAGAACACA 1046
            ||||||||||||||||||||||||
Db   1042 CAACCACACTTTCTCTAGAACACA 1067

RESULT 9
AAV74703
ID   AAV74703 standard; DNA; 2424 BP.
XX
AC   AAV74703;
XX
DT   16-MAR-1999 (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #392.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   1141..1200
FT   /*tag= a
FT   /note= "these bases represent a line of missing text in
FT   the sequence listing in the specification. They are
FT   included to maintain the nucleotide numbering given in
FT   the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997; 97EP-00100117.
XX
PR   05-JAN-1996; 96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX

```
CC   immunisation. The present sequence represents a partial nucleotide
CC   sequence (MurC ORF) of the MurC gene
XX
SQ   Sequence 660 BP; 233 A; 83 C; 120 G; 224 T; 0 U; 0 Other;

     Query Match             25.0%;  Score 330;  DB 2;  Length 660;
     Best Local Similarity   99.7%;  Pred. No. 2.3e-142;
     Matches 380; Conservative   0; Mismatches  1; Indels   0; Gaps   0;

Qy    666 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGC 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGC 60

Qy    726 TTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACATATGGTGA 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACATATGGTGA 120

Qy    786 CCATACAGTTTAAATGCTAATTGCGATTAGTTAGTTATTAGAGAAGCTAGATGT 845
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCATACAGTTTAAATGCTAATTGCGATTAGTTAGTTATTAGAGAAGCTAGATGT 180

Qy    846 TACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAATGAAAC 905
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAATGAAAC 240

Qy    906 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 965
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 300

Qy    966 TACAAATTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTCAACC 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TACAAATTGACACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTCAACC 360

Qy   1026 ACACACTTTCTCTAGAACACA 1046
          |||||||||||||||||||||
Db    361 ACACACTTTCTCTAGAACACA 381
```

RESULT 11
```
AAV53479/c
ID   AAV53479 standard; DNA; 619 BP.
XX
AC   AAV53479;
XX
DT   30-OCT-1998  (first entry)
XX
DE   DNA encoding a Staphylococcus aureus protein of unknown function.
XX
KW   Staphylococcus aureus protein; immune response induction; eye infection;
KW   antibody production; T-cell immune response; gastrointestinal infection;
KW   respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW   central nervous system; kidney infection; urinary tract infection;
KW   antimicrobial compound identification; broad spectrum antibiotic;
KW   therapy; ss.
XX
OS   Staphylococcus aureus.
XX
PN   EP841394-A2.
XX
PD   13-MAY-1998.
XX
PF   24-SEP-1997;  97EP-00307485.
XX
PR   24-SEP-1996;  96US-0027032P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Black MT, Hodgson JE, Knowles DJC,  Reichard RW, Nicholas RO;
PI   Burnham MKR,  Pratt JM,  Rosenberg M,  Ward JM,  Lonetto MA;
XX
DR   WPI; 1998-252940/23.
DR   P-PSDB; AAW77686.
```

```
Qy   1175 GTGCATCGTTAATTAATGAGATTCTATTAATGTATTAGAACAATTGATAATGCTGTTA 1234
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GTGCATCGTTAATTAATGAGATTCTATTAATGTATTAGAACAATTGATAATGCTGTTA 1860

Qy   1235 TTTTATTTATGGGTGCAGGTGATATTCAAAAATTACAAAATGCATATTTAGATAAATTAG 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTTTATTTATGGGTGCAGGTGATATTCAAAAATTACAAAATGCATATTTAGATAAATTAG 1920

Qy   1295 GCATGAAAAATGCGTTTTAA 1314
          ||||||||||||||||||||
Db   1921 GCATGAAAAATGCGTTTTAA 1940
```

RESULT 10
```
AAV80065
ID   AAV80065 standard; DNA; 660 BP.
XX
AC   AAV80065;
XX
DT   17-MAR-1999  (first entry)
XX
DE   Partial nucleotide sequence of the MurC gene.
XX
KW   MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW   bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW   immunogen; drug; genetic immunisation; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..660
FT                   /tag= a
FT                   /product= "partial MurC polypeptide"
XX
PN   EP889123-A2.
XX
PD   07-JAN-1999.
XX
PF   26-JUN-1998;  98EP-00305064.
XX
PR   03-JUL-1997;  97US-0052720P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Wallis NG,  Burnham MKR;
XX
DR   WPI; 1999-062655/06.
DR   P-PSDB; AAW89199.
XX
PT   New isolated MurC polypeptide from Staphylococcus aureus and related
PT   nucleic acid - useful in diagnosis, treatment and prevention of bacterial
PT   infections.
XX
PS   Claim 2; Page 4-5; 39pp; English.
XX
CC   The invention relates to a UDP-N-acetylmuramate:L-alanine ligase (MurC
CC   polypeptide) encoded by the S. aureus MurC gene. Host cells containing an
CC   expression system comprising the MurC gene can be used for the
CC   recombinant production of the polypeptide. Agonists or the MurC
CC   polypeptide are used to treat conditions requiring increased activity or
CC   expression of the polypeptide. Antagonists, inhibitory nucleic acid or
CC   competitive polypeptide are useful for inhibiting the polypeptide e.g.
CC   bacterial (especially S. aureus) infections. They are also useful against
CC   Helicobacter pylori infections and related cancers, ulcers and gastritis.
CC   The antibacterial agents are useful to treat in-dwelling devices for
CC   infection prevention or generally as wound treatments to prevent adhesion
CC   of bacteria to matrix proteins. The MurC polypeptide is also useful for
CC   diagnosing or prognosing a (susceptibility to) disease, for raising
CC   antibodies; to identify modulators or specific receptors; in rational
CC   drug design and as an immunogen for vaccines. The MurC gene sequences are
CC   useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
CC   for chromosomal mapping; to determine bacterial serotype; and for genetic
```

XX
PT New nucleic acid sequences from Staphylococcus aureus WCHU29 - useful in
PT vaccines and for treatment of bacterial infections of e.g. respiratory
PT tract and central nervous system.
XX
PS Claim 1; Page 146; 390pp; English.
XX
CC This sequence encodes a Staphylococcus aureus protein of unknown
CC function, and represents a DNA sequence of the invention. The DNA
CC sequences were isolated from Staphylococcus aureus WCHU29 (NCIMB 40771).
CC Host cells containing the DNA sequences are used to produce polypeptides
CC or fragments. The proteins are used in the treatment of disease, for
CC inducing an immune response by administering them, to produce antibody
CC and/or T-cell immune response. Antagonists of the proteins are used for
CC the inhibition of bacterial polypeptides. Conditions which may be treated
CC include bacterial infections, especially respiratory, cardiac,
CC gastrointestinal, central nervous, eye, kidney, urinary tract, skin,
CC bones and joints. The proteins can also be used to identify antimicrobial
CC compounds which are broad spectrum antibiotics, especially useful in the
CC treatment of H. pylori infection
XX
SQ Sequence 619 BP; 208 A; 117 C; 79 G; 215 T; 0 U; 0 Other;

Query Match 20.9%; Score 275; DB 2; Length 619;
Best Local Similarity 99.7%; Pred. No. 6.3e-117;
Matches 325; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy  721 ACTGCTTTGATGTGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
        ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  611 ACTGCTTTGATGTGATGTGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 552

Qy  781 GGTGACCACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTAGTTATTAGAGAAGCTA 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  551 GGTGACCACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTAGTTATTAGAGAAGCTA 492

Qy  841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTAAACGTCGTTCAAT 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  491 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTAAACGTCGTTCAAT 432

Qy  901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
        |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  431 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCACCATCCAAGAGAAATT 372

Qy  961 AGTGCTACAATTGAAACAGCAGCACGAAAGCACGAAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
        ||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  371 AGTGCTACAATTGACACAGCAGCACGACGAAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT 312

Qy 1021 CAACCACACACTTCTCTAGAACACA 1046
        |||||||||||||||||||||||||
Db  311 CAACCACACACTTCTCTAGAACACA 286
```

RESULT 12
AAS50351/c
ID AAS50351 standard; DNA; 257 BP.
XX
AC AAS50351;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1575.
XX
KW Antisense; ss; prokaryotic cellular proliferation; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX

PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 1; SEQ ID NO 2928; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence is an antisense oligonucleotide of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 257 BP; 84 A; 50 C; 31 G; 92 T; 0 U; 0 Other;

Query Match 19.5%; Score 257; DB 4; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.4e-108;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 153 ACCATTGATGCTAATAACATAAAGAAGAATATGGTAGTTATCAACGTAATGCATTCGC 212
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 257 ACCATTGATGCTAATAACATAAAGAAGAATATGGTAGTTATCAACGTAATGCATTCGC 198

Qy 213 GAGTAGCCATGAAGAAATAGTACGTGCACATCAATTGAAATTAGATGTTGTAAGTTATAA 272
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 197 GAGTAGCCATGAAGAAATAGTACGTGCACATCAATTGAAATTAGATGTTGTAAGTTATAA 138

Qy 273 TGATTTTTTAGGACAGATTATTGATCAATATACTTCAGTAGCTGTAACTGGTGCACATGG 332
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 TGATTTTTTAGGACAGATTATTGATCAATATACTTCAGTAGCTGTAACTGGTGCACATGG 78

Qy 333 TAAAACTTCTACAACAGGTTTATTATTATCACATGTTATGAATGGTGATAAAAAGACTTCATT 392
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 77 TAAAACTTCTACAACAGGTTTATTATTATCACATGTTATGAATGGTGATAAAAAGACTTCATT 18

Qy 393 TTTAATTGGTGATGGCA 409
       |||||||||||||||||
Db 17 TTTAATTGGTGATGGCA 1
```

RESULT 13
ACA17633/c
ID ACA17633 standard; DNA; 257 BP.
XX
AC ACA17633;
XX
DT 27-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)

Db   257 ACCATTTGATGCTAATAACATAAAGAAGATATGTGGTAGTTATACAGGTAATGCATTCGC 198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   213 GAGTAGCCATGAAGAAATAGTACGTGCACATCAATCAATTGAAATTAGATGTTGTAAGTTATAA 272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   197 GAGTAGCCATGAAGAAATAGTACGTGCACATCAATCAATTGAAATTAGATGTTGTAAGTTATATAA 138

Qy   273 TGATTTTTTAGGACAGATTATTGATCAATATACTTCAGTAGCTGTCAACTGCTGCACATGG 332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   137 TGATTTTTTAGGACAGATTATTGATCAATATACTTCAGTAGCTGTCAACTGCTGCACATGG 78

Qy   333 TAAAACTTCTACAACAGGTTTATTATTACACATGTTATGATGGTGTGATAAAAAGACTTCATT 392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   77 TAAAACTTCTACAACAGGTTTATTATTACACATGTTATGATGGTGTGATAAAAAGACTTCATT 18

Qy   393 TTTAATTGGTGATGGCA 409
          |||||||||||||||||

Db   17 TTTAATTGGTGATGGCA 1

RESULT 14
AAS49022/c
ID   AAS49022 standard; DNA; 225 BP.
XX
AC   AAS49022;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #246.
XX
KW   Antisense; ss; prokaryotic cellular proliferation; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,   Ohlsen KL,   Zyskind JW,   Wall D,   Trawick JD,   Carr GJ;
PI   Yamamoto RT,   Xu HH;
XX
DR   WPI; 2004-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 1; SEQ ID NO 1599; 51lpp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous

XX
DE   Prokaryotic essential gene antisense oligonucleotide #5503.
XX
KW   Antisense; ss; prokaryotic essential gene; cell proliferation;
KW   drug design.
XX
OS   Archaea.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-0072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 1; SEQ ID NO 5503; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the 6213
CC   antisense sequences of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 257 BP; 84 A; 50 C; 31 G; 92 T; 0 U; 0 Other;

Query Match   19.5%;   Score 257;   DB 7;   Length 257;
Best Local Similarity 100.0%;   Pred. No. 1.4e-108;
Matches   257; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   153 ACCATTTGATGCTAATAACATAAAGAAGATATGGTAGTTATACAAGGTAATGCATTCGC 212

Left column

```
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence is an antisense oligonucleotide of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 225 BP; 79 A; 44 C; 31 G; 71 T; 0 U; 0 Other;

Query Match            17.1%;   Score 225; DB 4;  Length 225;
Best Local Similarity 100.0%;   Pred. No. 8.8e-94;
Matches 225;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  299 AATATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGGTTTATTAT 358
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  225 AATATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGGTTTATTAT 166

Qy  359 CACATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGG 418
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165 CACATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGG 106

Qy  419 GATTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAA 478
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  105 GATTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAA 46

Qy  479 GTTATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATC 523
        |||||||||||||||||||||||||||||||||||||||||||||
Db   45 GTTATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATC 1

RESULT 15
AAS49054/c
ID   AAS49054 standard; DNA; 225 BP.
XX
AC   AAS49054;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #278.
XX
KW   Antisense; 88; prokaryotic cellular proliferation; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM ING
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 1; SEQ ID NO 1631; 51pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
```

Right column

```
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence is an antisense oligonucleotide of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 225 BP; 79 A; 44 C; 31 G; 71 T; 0 U; 0 Other;

Query Match            17.1%;   Score 225; DB 4;  Length 225;
Best Local Similarity 100.0%;   Pred. No. 8.8e-94;
Matches 225;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  299 AATATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGGTTTATTAT 358
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  225 AATATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGGTTTATTAT 166

Qy  359 CACATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGG 418
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165 CACATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGG 106

Qy  419 GATTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAA 478
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  105 GATTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAA 46

Qy  479 GTTATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATC 523
        |||||||||||||||||||||||||||||||||||||||||||||
Db   45 GTTATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATC 1

Search completed: June  6, 2004, 04:34:35
Job time : 639 secs
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       June  5, 2004, 21:59:24 ; Search time 636 Seconds
                              (without alignments)
                        8803.656 Million cell updates/sec

Title:        US-10-712-713-1
Perfect score: 1318
Sequence:     1 atgacacactcatttgt.........gaaaatgcgtttaagctt 1318
Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_29Jan04:*
               1: geneseqn1980s:*
               2: geneseqn1990s:*
               3: geneseqn2000s:*
               4: geneseqn2001as:*
               5: geneseqn2001bs:*
               6: geneseqn2002s:*
               7: geneseqn2003as:*
               8: geneseqn2003bs:*
               9: geneseqn2003cs:*
              10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1318	100.0	1318	4	AAS00189	Aas00189 S. aureus
2	1314	99.7	1335	9	ADD15462	Add15462 Staphyloc
3	1312.4	99.6	1335	9	ADD15464	Add15464 Staphyloc
4	1311	99.5	1332	7	ACF72770	Acf72770 Staphyloc
5	1310.8	99.5	1314	7	ACA19726	Aca19726 Prokaryot
6	1302.8	98.8	1335	4	AAS54558	Aas54558 Staphyloc
7	1299.8	98.6	1311	4	AAS51660	Aas51660 Staphyloc
8	1293.2	98.1	1351	2	AAV99650	Aav99650 UDP-N-ace
9	1254	95.1	2424	2	AAV74703	Aav74703 Staphyloc
10	922	70.0	1329	6	ABN92833	Abn92833 Staphyloc
11	915.8	69.5	1311	7	ACA47124	Aca47124 Prokaryot
12	633	48.0	660	2	AAV80065	Aav80065 Partial n
13	597.6	45.3	1308	7	ACA22277	Aca22277 Prokaryot
c 14	570.4	43.3	619	2	AAV53479	Aav53479 DNA encod
15	555	42.1	1371	4	AAS53136	Aas53136 Enterococ
16	553.4	42.0	1335	7	ACA33252	Aca33252 Prokaryot
c 17	523	39.7	110000	6	ABQ69245_16	Continuation (17 o
18	517.4	39.3	110000	6	ABQ67195_0	Abq67195 Listeria
19	517.4	39.3	110000	6	ABQ67195_1	Continuation (2 of
20	516.4	39.2	1344	7	ACA36934	Aca36934 Prokaryot
c 21	516.4	39.2	110000	6	ABA03041_16	Continuation (17 o
22	515	39.1	1332	7	ACA33597	Aca33597 Prokaryot
c 23	494.8	37.5	738	7	ACA47261	Aca47261 Prokaryot

c 24	456	34.6	110000	6	ABA90521_21	Continuation (22 o
25	450.8	34.2	1329	6	ABN68449	Abn68449 Streptoco
c 26	450.8	34.2	110000	6	ABN71527_15	Continuation (16 o
27	447.8	34.0	1326	6	ABN68450	Abn68450 Streptoco
28	447.8	34.0	1329	7	ACA50341	Aca50341 Prokaryot
29	426.2	32.3	1329	7	ACA48125	Aca48125 Prokaryot
30	417	31.6	1335	8	ADB11619	Adb11619 Alloiococ
c 31	417	31.6	110000	8	ADB12064_14	Continuation (15 o
32	402.4	30.5	1335	7	AAS55772	Aas55772 Streptoco
33	402.4	30.5	1335	7	ACA50019	Aca50019 Prokaryot
34	400.8	30.4	1332	7	ABX07268	Abx07268 S. pneumo
35	400.8	30.4	11864	2	AAV52194	Aav52194 Streptoco
c 36	400.8	30.4	110000	7	ABS56454_13	Continuation (14 o
37	389.6	29.6	888	3	AAA26910	Aaa26910 Essential
38	389.6	29.6	888	4	AAF91592	Aaf91592 Staphyloc
39	389.6	29.6	888	4	AAS08061	Aas08061 Staphyloc
40	389.6	29.6	888	9	ADD67439	Add67439 Antibacte
41	389.6	29.6	888	9	ADE73663	Ade73663 Mutant ba
42	370.2	28.1	1267	2	AAV27381	Aav27381 Streptoco
43	370.2	28.1	1267	6	ABQ84849	Abq84849 S. pneumo
44	370.2	28.1	1267	9	ADC45196	Adc45196 S. pneumo
45	354.4	26.9	873	9	ADC93633	Adc93633 E. faeciu

ALIGNMENTS

```
RESULT 1
AAS00189
ID   AAS00189 standard; DNA; 1318 BP.
XX
AC   AAS00189;
XX
DT   04-JUL-2001  (first entry)
XX
DE   S. aureus DNA encoding UDP-N-acetylmuramate:L-alanine ligase, MurC.
XX
KW   UDP-N-acetylmuramate:L-alanine ligase; MurC; immunogen; vaccine;
KW   antibody; wound infection; cellulitis; burn infection; eyelid infection;
KW   food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW   skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW   Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW   ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1314
FT                   /*tag= a
FT                   /product= "MurC"
XX
PN   WO200116292-A2.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US023773.
XX
PR   01-SEP-1999; 99US-0151933P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Choi GH;
XX
DR   WPI; 2001-183259/18.
DR   P-PSDB; AAU00828.
XX
PT   New isolated nucleic acid for use in diagnosing Staphylococcus infections
PT   and in vaccines for eliciting immune responses to the infections.
XX
PS   Claim 1; Page 14; 225pp; English.
XX
CC   The sequence encodes S. aureus MurC (UDP-N-acetylmuramate:L-alanine
CC   ligase). The polynucleotides of the invention are used to detect
```

CC   Staphylococcus nucleic acids in a biological sample from an animal for
CC   diagnosing Staphylococcus infections. The polypeptides of the invention
CC   are used to detect anti-Staphylococcus antibodies in a biological sample
CC   from an animal to diagnose Staphylococcus infections. The polypeptides
CC   are also used in vaccines to elicit protective antibodies in an animal to
CC   a member of the Staphylococcus genus and for preventing or attenuating an
CC   infection caused by a member of the Staphylococcus genus e.g wound
CC   infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC   joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC   scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC   disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC   polynucleotides may be used in vaccines and for preventing or
CC   attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC   may be used to purify, detect and target the polypeptides in in vitro and
CC   in vivo diagnostic and therapeutic methods
XX
SQ   Sequence 1318 BP; 464 A; 162 C; 251 G; 441 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1318; DB 4;   Length 1318;
Best Local Similarity   100.0%;   Pred. No. 2.7e-249;
Matches 1318;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGACACACATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGACACACATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60

Qy   61 ATGCATGATTAGGACATGAGAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATGCATGATTAGGACATGAGAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120

Qy   121 GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAGAAATAGTACGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAGAAATAGTACGTGCA 240

Qy   241 CATCAATTGAAATTGAGATGTGTGTAAGTTATAATGATTTTTTAGGACAGAGATTATTGATCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCAATTGAAATTGAGATGTGTGTAAGTTATAATGATTTTTTAGGACAGAGATTATTGATCAA 300

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACAGGTTATTATCA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACAGGTTATTATCA 360

Qy   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATTATGACAAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATTATGACAAAATATTGATTTCGATCATCCTGATTATTTTAAA 540

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAAGGTATT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAAGGTATT 600

Qy   601 ATTGCTTGGGGTGGATGATGAACATCGAAGCAGGATGTTCCAATTATTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGCTTGGGGTGGATGATGAACATCGAAGCAGGATGTTCCAATTATTATTAT 660

Qy   661 TATGGATTAAAGATTCGGATGACATTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTAAAGATTCGGATGACATTATGCTCAAAATATTCAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCGTGTCTCCACAATAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCGTGTCTCCACAATAT 780

Qy   781 GGTGACCATACAGTTTAAATGCATTAGAGCTGTAATTGCCGATTAGTAGTTATTAGAGAAGCTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGTGACCATACAGTTTAAATGCATTAGAGCTGTAATTGCCGATTAGTAGTTATTAGAGAAGCTA 840

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAAT 900

Qy   901 GAAACTACAATGCAAATCAAGTTAATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAACTACAATGCAAATCAAGTTAATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Qy   961 AGTGCTACAATTGAAACAGCACGACGAAGATATCCACATAAGAAGAAGTTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGTGCTACAATTGAAACAGCACGACGAAGATATCCACATAAGAAGAAGTTGTTGCAGTATTT 1020

Qy   1021 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTAAGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTAAGT 1080

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140

Qy   1141 TTAACGATACAAGATTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTAATTAAGATTCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTAACGATACAAGATTAATTGATAAAATTGATAAAATTGAAGGTGCATCGTAATTAAGATTCT 1200

Qy   1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTATTTATTGGGTGCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTATTTATTGGGTGCAGGTGATATT 1260

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTTAAGCTT 1318
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTTAAGCTT 1318

RESULT 2
ADD15462
ID   ADD15462 standard; DNA; 1335 BP.
XX
AC   ADD15462;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Staphylococcus aureus murC DNA (SeqID 27).
XX
KW   microbial; antimicrobial; membrane biosynthesis; pathogenic;
KW   immunological response; vaccination; surface disinfectant;
KW   personal hygiene application; food preservative; ds; gene; murC;
KW   UDP-N-acetylmuramate-alanine ligase.
XX
OS   Staphylococcus aureus.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1335
FT   /*tag= a
FT   /product= "MurC protein"
XX
PN   WO2003025007-A2.
XX
PD   27-MAR-2003.
XX
PF   20-SEP-2002; 2002WO-CA001428.
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.

```
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 27; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polynucleotide
CC   sequence is DNA predicted from the genomic sequence of S. aureus UDP-N-
CC   acetylmuramate-alanine ligase (murC) of the invention.
XX
SQ   Sequence 1335 BP; 472 A; 161 C; 255 G; 447 T; 0 U; 0 Other;

Query Match             99.7%;  Score 1314;  DB 9;  Length 1335;
Best Local Similarity  100.0%;  Pred. No. 1.6e-248;
Matches 1314;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGACACACTATCATTTGTCGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATTTGTCGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 81

Qy     61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 141

Qy    121 GTTGCTCCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GTTGCTCCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 201

Qy    181 GATAATGGTAGTTATACAAGGTAATGCATTCGGCAGTAATGATTTTTAGGACAGATTATTGATCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GATAATGGTAGTTATACAAGGTAATGCATTCGGCAGTAATGATTTTTAGGACAGATTATTGATCAA 261

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 321

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTCACATGGTAAAACTTCTACAACAGTTTATTATTCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTCACATGGTAAAACTTCTACAACAGTTTATTATTCA 381

Qy    361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441
```

```
Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 501

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 561

Qy    541 GATATTAATGATGTTTTTGATGCATTCCAAGAAAATGGCACATAAATGTAAAAAAGGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTTGATGCATTCCAAGAAAATGGCACATAAATGTAAAAAAGGTATT 621

Qy    601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTAT 681

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATATTCAAATTACGGATAAAGGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATATTCAAATTACGGATAAAGGT 741

Qy    721 ACTGCTTTTGATGTGATGTGGAGTTTATGATCCGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTTGATGTGATGTGGAGTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTGTAATTGCGATTAGTTATTTAGAGAAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTAAATGCATTAGCTGTGTAATTGCGATTAGTTATTTAGAGAAGCTA 861

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAAACGTCGTTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAAACGTCGTTTCAAT 921

Qy    901 GAAACTACAATTGCAAATCAAGTTAGATGATTATGCACACCATCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAATTGCAAATCAAGTTAGATGATTATGCACACCATCCAAGAGAAATT 981

Qy    961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTGTGTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTGTGTGCAGTATTT 1041

Qy   1021 CAACCACACACTTTCTCTAGAACACAGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCTAGAACACAGCATTTTTAAATGAATTGCAGAAAGTTTAAGT 1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1161

Qy   1141 TTAACGATACAAGATTTAATTGATAAAATTGAGAGGTGCATCGTTAATTAATGAAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTTAATTGATAAAATTGAGAGGTGCATCGTTAATTAATGAAGATTCT 1221

Qy   1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTTATTTATTTATTTATGGGTCAGTGATATT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGTATTAGAACAATTTGATAATGCTGTTTATTTATTTATTTATGGGTGCAGTGATATT 1281

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGCATGAAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAAATGCATATTTAGATAAATTAGCATGAAAAATGCGTTTTAA 1335
```

```
KW   personal hygiene application; food preservative; ds; gene; murC;
KW   UDP-N-acetylmuramate-alanine ligase.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1335
FT                   /*tag= a
FT                   /product= "MurC protein"
XX
PN   WO200325007-A2.
XX
PD   27-MAR-2003.
XX
PF   20-SEP-2002; 2002WO-CA001428.
XX
PR   21-SEP-2001; 2001US-0323992P.
PR   21-SEP-2001; 2001US-0324152P.
PR   25-SEP-2001; 2001US-0324692P.
PR   26-OCT-2001; 2001US-0339924P.
PR   29-OCT-2001; 2001US-0350973P.
PR   30-OCT-2001; 2001US-0340924P.
PR   27-NOV-2001; 2001US-0333666P.
PR   18-DEC-2001; 2001US-0341732P.
PR   18-DEC-2001; 2001US-0341776P.
PR   19-DEC-2001; 2001US-0341949P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A,   Dharamsi A,   Vedadi M,   Alam MZ,   Awrey D,   Beattie B;
PI   Domagala M,   Houston S,   Kanagarajah D,   Nethery K,   Ng I,   Mansoury K;
PI   Mcdonald M,   Pinder B,   Viola C,   Wrezel O;
XX
DR   WPI; 2003-468119/44.
DR   P-PSDB; ADD15465.
XX
PT   Novel crystallized recombinant polypeptides from Staphylococcus aureus,
PT   Streptococcus pneumoniae and Escherichia coli and which are involved in
PT   membrane biosynthesis, useful as targets for pathogenic bacteria.
XX
PS   Claim 20; SEQ ID NO 29; 325pp; English.
XX
CC   This invention relates to the structural and functional characterisation
CC   of microbial polypeptides from Staphylococcus aureus (S. aureus),
CC   Streptococcus pneumoniae (S. pneumoniae) and Escherichia coli (E. coli)
CC   that provide novel antimicrobial targets. Specifically, it refers to
CC   polypeptides that are involved in membrane biosynthesis, which play a
CC   critical role in the life cycle and viability of their pathogenic species
CC   of origin, and hence provide valuable drug targets. Furthermore, the
CC   invention describes modified version of these proteins that facilitate
CC   characterisation by labelling with isotopic or heavy atoms, and also
CC   fusion proteins. These proteins provide structural and functional
CC   information to aid the discovery of therapeutic molecules to treat
CC   disorders associated with a particular pathogenic species. As such, they
CC   are useful for inducing an immunological response in an individual and as
CC   an antigen for vaccination purposes. The polypeptides are also useful for
CC   developing antimicrobial agents for use as surface disinfectants,
CC   personal hygiene applications and as food preservatives or in treating
CC   food products to eliminate potential pathogens. This polynucleotide
CC   sequence is the experimentally predicted DNA of S. aureus UDP-N-
CC   acetylmuramate-alanine ligase (murC) of the invention.
XX
SQ   Sequence 1335 BP; 472 A; 160 C; 255 G; 448 T; 0 U; 0 Other;

Query Match             99.6%;  Score 1312.4;  DB 9;  Length 1335;
Best Local Similarity   99.9%;  Pred. No. 3.4e-248;
Matches 1313;   Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGACACACTATCATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 81

Qy     61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
```

Sequence alignment block (Db / Qy rows):

```
Db    82 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 141
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 201
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAAGAAATAGTACGTGCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAGAAGAAATAGTACGTGCA 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 CATCAATTGAAATTAGAGTGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CATCAATTGAAATTAGAGTGTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 321
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACAGGTTTATTATCA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACAGGTTTATTATCA 381
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGTATGGGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGTATGGGA 441
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTGAGGCATGTGAAATATAGACGTCACTTTTAAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 TTGCCTGAAAGTGATTATTTCGCTTTGAGGCATGTGAAATATAGACGTCACTTTTAAGT 501
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 TATAAACCTGATTACGCAATTATGACAATATTGATTTCGATCATCCTGATTATTTTAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 TATAAACCTGATTACGCAATTATGACAATATTGATTTCGATCATCCTGATTATTTTAAA 561
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 621
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 ATTGCTTGGGGTGATGAATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ATTGCTTGGGGTGATGAATGAACATTTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 681
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 TATGGATTTAAAGATTCGGATGACATTCAAATTATTCAAAATATTCAAATTACGGATAAAGGT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAATTACGGATAAAGGT 741
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 ACTGCTTTGATGTGTATGGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACTGCTTTGATGTGTATGGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 801
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTATTTAGAGAGCTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTATTTAGAGAGCTA 861
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 921
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 981
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 AGTGCTACAAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGTGCTACAAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021 CAACCACACTTTCTCTAGAACAACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1042 CAACCACACTTTCTCTAGAACAACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 1101
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1161
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTCATCGTTAATTAATGAAGATTCT 1200
```

Db   142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 ATTAATGTATTAGAACAATTGATAAAAATGAAGGTGCATCGTTAATTAATGAAGATTCT 1221

Db   181 GATATGGTAGTTATACAAGGTAATCATTCGGCAGTAGCCATGAGGAATAGTACGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   202 GATATGGTAGTTATACAAGGTAATCATTCGGCAGTAGCCATGAGGAATAGTACGTGGGA 261

Db   241 CATCAATTGAAATTAGATGTTGTAAGTTAATAATGATTTTTTAGGACAGATTATTGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   262 CATCAATTGAAATTAGATGTTGTAAGTTAATAATGATTTTTTAGGACAGATTATTGATCAA 321

Db   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTGTAAAACTTCTACACACAGGTTATTATCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTGTAAAACTTCTACAACAGGTTTATTATCA 381

Db   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 441

Db   421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGCATGTGAATATAGACGTCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   442 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGCATGTGAATATAGACGTCACTTTTTAAGT 501

Db   481 TATAAACCTGATTACGCAATTATGACAAATTATTGATTCGATCATCCTGATTATTTTAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   502 TATAAACCTGATTACGCAATTATGACAAATTATTGATTCGATCATCCTGATTATTTTAAA 561

Db   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGTATT 621

Db   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTTCCAATTTATTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTTCCAATTTATTAT 681

Db   661 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   682 TATGGATTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 741

Db   721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 801

Db   781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATGCGTAATTAGGAAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   802 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATGCGTAATTAGTTAGGAAGCTA 861

Db   841 GATGTTACAAATATTAAAGAAGAGCATTAAGCATTGGTGTGGTGTAAACGTTCGTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   862 GATGTTACAAATATTAAAGAAGAGCATTAAGCATTGGTGTGGTGTAAACGTCGTTCAAT 921

Db   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 981

Db   961 AGTGCTACAACAGCAGCGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   982 AGTGCTACAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Db   1021 CAACCACACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAGTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1042 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTAAGT 1101

Db   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTCGGCGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTCGGCGCA 1161

Db   1141 TTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1162 TTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1221

Db   1201 ATTAATGTATTAGAACAATTGATAAATGTGCTAATTGATAAAAATTAGGCATGAAAAATGCGTTTTAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1222 ATTAATGTATTAGAACAATTGATAAATGTGCTAATTGATAAAAATTAGGCATGAAAAATGCGTTTTAA 1281

Db   1261 CAAAAAATTACAAAATGCATATTTAGATAAATGATAAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||||
Qy   1282 CAAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTAA 1335

RESULT 4

ACF72770

ID   ACF72770 standard; DNA; 1332 BP.
XX
AC   ACF72770;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Staphylococcus aureus DNA #450.
XX
KW   Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW   enzymatic assay; antibiotic target; gene; ds.
XX
OS   Staphylococcus aureus.
XX
PN   WO200294868-A2.
XX
PD   28-NOV-2002.
XX
PF   27-MAR-2002; 2002WO-IB002637.
XX
PR   27-MAR-2001; 2001GB-00007661.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V, Mora M, Scarselli M;
XX
DR   WPI; 2003-120786/11.
DR   P-PSDB; ABM71210.
XX
PT   New Staphylococcus aureus protein, useful as a vaccine for treating or
PT   preventing Staphylococcal infection, specifically an infection caused by
PT   S. aureus, e.g. sepsis.
XX
PS   Claim 6; SEQ ID NO 899; 49pp; English.
XX
CC   The invention relates to novel genes and encoded proteins from
CC   Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC   nucleic acid encoding the protein, or an antibody to the protein, is
CC   useful as a pharmaceutical, particularly as a vaccine for treating or
CC   preventing infection due to Staphylococcus bacteria, specifically an
CC   infection caused by S. aureus. The composition is particularly useful for
CC   treating or preventing sepsis in a patient. The composition can also be
CC   used for diagnostics. The protein is also used in an assay for enzymatic
CC   studies and as a target for antibiotics. This sequence represents one of
CC   the novel S. aureus genes of the invention
XX
SQ   Sequence 1332 BP; 470 A; 161 C; 255 G; 446 T; 0 U; 0 Other;

Query Match   99.5%; Score 1311; DB 7; Length 1332;
Best Local Similarity   100.0%; Pred. No. 6.4e-248;
Matches 1311; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACACTACTATCATTTGTCGGAATTAAAGTTCTGCCATGAGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACTACTATCATTTGTCGGAATTAAAGTTCTGCCATGAGTTCATTAGCACAAATC 81

Qy   61 ATGCATGATGATTAGGACATGAAGATGAAGAGTTCAAGGATCGAAGATCGGATATTGAAGATTCTTACAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATGATTAGGACATGAAGATGAAGAGTTCAAGGATCGAAGATCGGATATTGAAGATTCTTACAGAA 141

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180

```
Db   1222 ATTAATGTATTAGAACAATTTGATAATGCTGTATTTATTTATTGGGTGCAGGTGATATT 1281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 CAAAAATTACAAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTT 1311
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTT 1332


RESULT 5
ACA19726
ID   ACA19726 standard; DNA; 1314 BP.
XX
AC   ACA19726;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #1383.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus aureus.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU15856.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 7596; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts;(9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising screening in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential gene. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1314 BP; 464 A; 161 C; 250 G; 439 T; 0 U; 0 Other;

Query Match             99.5%;  Score 1310.8;  DB 7;  Length 1314;
Best Local Similarity   99.8%;  Pred. No. 7e-248;
Matches 1312;  Mismatches 0;  Conservative 0;  Indels 2;  Gaps 0;

Qy    1 ATGACACCATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGACACCATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60

Qy   61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTAGTATTTACAGAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTAGTATTTACAGAA 120

Qy  121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180

Qy  181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAAGAAATAGTACGTGCA 240

Qy  241 CATCAATTGAAATTAGTAGTGTTGTAAGTTATATGATTTTTAGGACAGATTATTGATCAA 300
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  241 CATCAATTGAAATTAGTAGTGTTGTAAGTTATATGATTTTTAGGACAGATTATTGATCAA 300

Qy  301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAACTTCACAACAGGTTTATTATCA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAACTTCACAACAGGTTTATTATCA 360

Qy  361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy  421 TTGCCTGAAAGTGATTGATTATTCGCTTTTGAGGCATGGAATATAGACGTCACTTTTAAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TTGCCTGAAAGTGATTGATTATTCGCTTTTGAGGCATGGAATATAGACGTCACTTTTAAGT 480

Qy  481 TATAAACCTGATTACGCCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TATAAACCTGATTACGCCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540

Qy  541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600

Qy  601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660

Qy  661 TATGGATTTAAAGATTCGGATGACATTGATCTCAAAATTACCGATAAAGGT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TATGGATTTAAAGATTCGGATGACATTGATCTCAAAATTACCGATAAAGGT 720

Qy  721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy  781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCTGTAATTAGAGAAGCTA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTAGCTGTAATTAGAGAAGCTA 840

Qy  841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT 900
```

```
Qy  901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT  960

Qy  961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020

Qy 1021 CAACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAGTTTAAGT     1080
        ||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db 1021 CAACCACACTTTCTCTAGAACACAAGCATTTTAAATGAATTTGCAGAAGTTTAAGT     1080

Qy 1081 AAAGCAGATCGTGATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGCGCA   1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 AAAGCAGATCGTGATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTGCGCA   1140

Qy 1141 TTAACGATACAAGATTCTTATGTGATAAAATTGAAAGGTGCATCGTTAATTAATGAAGATTCT 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200

Qy 1201 ATTAATGTTATTAGAACAATTTGATAATGCTGTGTATTTTATTATTATGGGTGCAGTGATATT 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 ATTAATGTTATTAGAACAATTTGATAATGCTGTGTATTTTATTATTATGGGTGCAGTGATATT 1260

Qy 1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAAATGCGTTTTAA 1314
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAAATGCGTTTTAA 1314
```

RESULT 6
AAS54558
ID   AAS54558 standard; DNA; 1335 BP.
XX
AC   AAS54558;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #870.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   27-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36699.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 8195; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,

---

CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1335 BP; 472 A; 164 C; 255 G; 444 T; 0 U; 0 Other;

Query Match   98.8%; Score 1302.8; DB 4; Length 1335;
Best Local Similarity 99.5%; Pred. No. 2.6e-246;
Matches 1307; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy    1 ATGACACACTATCATCATTTGTCGGAATTAAAGGTTCTGGCGATGAGTTCATTAGCACAAATC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATCATTTGTCGGAATTAAAGGTTCTGGCGATGAGTTCATTAGCACAAATC   81

Qy   61 ATGCATGATTTAGGACATGAAGTTCAAGATCGGAATTCAAGGTCAGATATTGAGAACTACGTATTACAGAA  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATTTAGGACATGAAGTTCAAGATCGGAATTCAAGGTCAGATATTGAGAACTACGTATTACAGAA  141

Qy  121 GTTGTCTCTTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  180
        ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  142 GTTGTCTCTTAGAAATAAGGGGATAAAAAATATTACCATTTGGTGCTAATAACATAAAAGAA  201

Qy  181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA  261

Qy  241 CATCAATTGAAATTAGATGTGTGAAGTATATAGATTTTTAGGACAGAGATTATTGATCAA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 CATCAATTGAAATTAGATGTGTGAAGTATATAGATTTTTAGGACAGAGATTATTGATCAA  321

Qy  301 TATACTTCAGTAGCTGTAACTCGGTGCACATGGTAAAACTTCTACAACAGGTTTTATTATCA  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  322 TATACTTCAGTAGCTGTAACTCGGTGCACATGGTAAAACTTCTACAACAGGTTTTATTATCA  381

Qy  361 CATGGTTATGGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGCGCACAGGTATGGGA  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  382 CATGGTTATGGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGCGCACAGGTATGGGA  441

Qy  421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGGCATGTGAATATAGACGTCACTTTTTAAGT  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGGCATGTGAATATAGACGTCACTTTTTAAGT  501

Qy  481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTCAA  561

Qy  541 GATATTAATGAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTAAAAAAGTATT  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  562 GATATTAATGAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTAAAAAAGTATT  621

Qy  601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  622 ATTGCTTGGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTTATTAC  681

Qy  661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT  741

Qy  721 ACTGCTTTGATTGTATGTGATGGATGCGAGTTTTATGATCACTTCCTGTCTCCACAATAT  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33801.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 4242; 51pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1311 BP; 462 A; 164 C; 250 G; 435 T; 0 U; 0 Other;

Query Match             98.6%;  Score 1299.8; DB 4;  Length 1311;
Best Local Similarity   99.5%;  Pred. No. 1e-245;
Matches 1304;  Conservative   0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         1 ATGACACCACTATCATTCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTGAGCACAAAATC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGACACCACTATCATTCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTGAGCACAAAATC 60

Qy        61 ATGCATGATTAGGACATGAGAGTTCAAGGATCGGATCGGATATTGAGAACTACGTATTTACAGAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATGCATGATTAGGACATGAGAGTTCAAGGATCGGATCGGATATTGAGAACTACGTATTTACAGAA 120

Qy       121 GTTGCTCTTAGAAATAAGGGGGATAAAAAAAATTACCATTGTGATGGCTAATACCATAAAAGAA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GTTGCTCTTAGAAATAAGGGGGATAAAAAAAATTACCATTGTGATGGCTAATACCATAAAAGAA 180

Qy       181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAGAAAAGTAGTACGTGCA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAGAAAAGTAGTACGTGCA 240

Qy       241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300

Qy       301 TATACTTCAGTAGCTGTAACTCGGTGCACATGGTAAAACTTCTACAACAGGTTTATATTCA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TATACTTCAGTAGCTGTAACTCGGTGCACATGGTAAAACTTCTACAACAGGTTTATATTCA 360

Qy       361 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CATGTTATGAATGGTGATAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy       421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480

Qy       481 TATAAACCTGATTACGCAATTATGCAATATATTCGATCATCCTGATTATTTTAAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TATAAACCTGATTACGCAATTATGCAATATATTCGATCATCCTGATTATTTTCAAA 540

Qy       541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
```

```
Db       742 ACTGCTCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 801
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTTATTTAGAGAAGCTA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       802 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTTATTTAGAGAAGCTA 861
             
Qy       841 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTTGGTGGTGTGTTAAACGTCGTTTCAAT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       862 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTTGGTGGTGTGTTAAACGTCGTTTCAAT 921

Qy       901 GAAACTACAATTGCAAATCAAGTTATTGTGATGATTATGCACACCATCCAAGAGAAATT 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       922 GAAACTACAATTGCAAATCAAGTTATCGTAGATGATTATGCACACCATCCAAGAGAAATT 981

Qy       961 AGTGTTACAAGTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGAGTTGTTGCAGTATTT 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       982 AGTGTTACAAGTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGAGTTGTTGCAGTATTT 1041

Qy      1021 CAACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1042 CAACCACACTTTCTCTAGAACACAAGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1101

Qy      1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAGAAAATACTGCGCA 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAGAAAATACTGCGCA 1161

Qy      1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1221

Qy      1201 ATTAATGATTAGAACAATTTGATAATGCTGTTATTTATTTTATTTATTTATGGGTGCAGGTGATATT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1222 ATTAATGATTAGAACAATTTGATAATGCTGTTATTTATTTTATTTATTTATGGGTGCAGGTGATATT 1281

Qy      1261 CAAAAATTACAAAATGCATATTAGATAATAATTAGGCATGAAAATGCGTTTTAA 1314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1282 CAAAAATTACAAAATGCATATTTAGATAATAATTAGGCATGAAAATGCGTTTTAA 1335
```

```
RESULT 7
AAS51660
ID   AAS51660 standard; DNA; 1311 BP.
XX
AC   AAS51660;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #77.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
```

```
XX
PD   07-JAN-1999.
XX
PF   26-JUN-1998;   98EP-00305064.
XX
PR   03-JUL-1997;   97US-0052720P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Wallis NG, Burnham MKR;
XX
DR   WPI; 1999-062655/06.
DR   P-PSDB; AAW87771.
XX
PT   New isolated MurC polypeptide from staphylococcus aureus and related
PT   nucleic acid - useful in diagnosis, treatment and prevention of bacterial
PT   infections.
XX
PS   Claim 2; Page 3-4; 39pp; English.
XX
CC   The present sequence represents a MurC gene encoding a Staphylococcus
CC   aureus UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide). Host
CC   cell containing an expression system comprising the MurC gene can be used
CC   for the recombinant production of the polypeptide. Agonists or the MurC
CC   polypeptide are used to treat conditions requiring increased activity or
CC   expression of the polypeptide. Antagonists, inhibitory nucleic acid or
CC   competitive polypeptide are useful for inhibiting the polypeptide e.g.
CC   bacterial (especially S. aureus) infections. They are also useful against
CC   Helicobacter pylori infections and related cancers, ulcers and gastritis.
CC   The antibacterial agents are useful to treat in-dwelling devices for
CC   infection prevention or generally as wound treatments to prevent adhesion
CC   of bacteria to matrix proteins. The MurC polypeptide is also useful for
CC   diagnosing or prognosing a (susceptibility to) disease, for raising
CC   antibodies; to identify modulators or specific receptors; in rational
CC   drug design and as an immunogen for vaccines. The MurC gene sequences are
CC   useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
CC   for chromosomal mapping; to determine bacterial serotype; and for genetic
CC   immunisation
XX
SQ   Sequence 1351 BP; 471 A; 166 C; 260 G; 454 T; 0 U; 0 Other;

Query Match       98.1%;   Score 1293.2;   DB 2;   Length 1351;
Best Local Similarity   99.0%;   Pred. No. 2e-244;
Matches 1301;   Conservative   0;   Mismatches   13;   Indels   0;   Gaps   0;

Qy       1 ATGACACACTATCATTTGTCCGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACACAATC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22 ATGACACACTATCATTTGTCCGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACACAATC   81

Qy      61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA   141

Qy     121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGGTGCTAATAACATAAAAGAA   201

Qy     181 GATATGGTAGTTATACAAGGTAATCATTGCGCGAGTAGCCATGGAGAAAATAGTACGTGCA   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     202 GATATGGTAGTTATACAAGGTAATCATTGCGCGAGTAGCCATGGAGAAAATAGTACGTGCA   261

Qy     241 CATCAATTGAAATTAGAATGTGTAAGTTGTAAGTTATAATGTAATAATATTGATCAA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     262 CATCAATTGAAATTAGAATGTGTAAGTTGTAAGTTATAATGTAATAATATTGATCAA   321

Qy     301 TATACTTCAGTAGCTGTCAACTGGTGCACATCGTAAAACTTCTACACACAGGTTTATTATCA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322 TATACTTCAGTAGCTGTCAACTGGTGCACATCGTAAAACTTCTACACACAGGTTTATTATCA   381

Qy     361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGCGCACAGGTATGGGA   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGCGCACAGGTATGGGA   441
```

```
Qy     601 ATTGCTCGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ATTGCTCGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAC   660

Qy     661 TATGGATTAAAGATCGGATGACACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TATGGATTAAAGATCGGATGACACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT   720

Qy     721 ACTGCTTTGATGTGTATGTGAGTTTATGATCACTTCCTGTCTCCACAATAT   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACTGCTTTGATGTGTATGTGGATGTGAGTTTATGATCACTTCCTGTCTCCACAATAT   780

Qy     781 GGTGACCATACAGTTTAAAATGCATTAGCTGTAATTGCCATTAGTTATTAGAGAAGCTA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GGTGACCATACAGTTTAAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGCTA   840

Qy     841 GATGTTACAAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTCAAT   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GATGTTACAAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT   900

Qy     901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAACTACAATTGCAAATCAAGTTATCGTAGATGATTATGCACACCATCCAAGAGAAATT   960

Qy     961 AGTGCTACAATTGAAACAGCCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AGTGCTACAATTGAAACAGCCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT   1020

Qy    1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTCGAGAAAGTTAAGT   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTCGAGAAAGTTTAAGT   1080

Qy    1081 AAAGCAGATCGTGATTCCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGCGCA   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGCGCA   1140

Qy    1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200

Qy    1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTATTTATGGGTGCAGGTGATATT   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTGTTTATTTATGGGTGCAGGTGATATT   1260

Qy    1261 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTT   1311
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CAAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAATGCGTTT   1311
```

```
RESULT 8
AAV99650
ID   AAV99650 standard; DNA; 1351 BP.
XX
AC   AAV99650;
XX
DT   17-MAR-1999 (first entry)
XX
DE   UDP-N-acetylmuramate:L-alanine ligase (MurC polypeptide) encoding DNA.
XX
KW   MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW   bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW   immunogen; drug; genetic immunisation; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key            Location/Qualifiers
FT   CDS            22..1335
FT                  /tag= a
FT                  /gene= "MurC"
FT                  /product= "UDP-N-acetylmuramate:L-alanine ligase (MurC
FT                  polypeptide)"
XX
PN   EP889123-A2.
```

```
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1141..1200
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                   the sequence listing in the specification. They are
FT                   included to maintain the nucleotide numbering given in
FT                   the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;  97EP-00100117.
XX
PR   05-JAN-1996;  96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA,  Choi GH,  Barash SC,  Dillon PJ,  Fannon MR,  Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 1287-1288; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 2424 BP; 839 A; 302 C; 462 G; 759 T; 0 U; 62 Other;

Query Match             95.1%;  Score 1254;  DB 2;  Length 2424;
Best Local Similarity   95.4%;  Pred. No. 1.1e-236;
Matches 1254;  Conservative   0;  Mismatches 60;  Indels   0;  Gaps   0;

Qy       1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 686

Qy      61 ATGCATGATTTAGGACATGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 120
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     687 ATGCATGATTTAGGACATGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 746

Qy     121 GTTGCTTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     747 GTTGCTTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 806

Qy     181 GATATGTGAGTTATACAACGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy     421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 501

Qy     481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTCAAA 561

Qy     541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 621

Qy     601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAC 681

Qy     661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 741

Qy     721 ACTGCTTTTCGATGTGTATGTGGATGCGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 ACTGCTTTTCGATGTGTATGTGGATGCGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy     781 GGTGACCATACAGTTTTAAATGCATTAAGTGCGATTAGTTATTTAGAGAAGCTA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     802 GGTGACCATACAGTTTTAAATGCATTAAGTGCGATTAGTTATTTAGAGAAGCTA 861

Qy     841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 921

Qy     901 GAAACTACACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 GAAACTACACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 981

Qy     961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 AGTGCTACAATTGACACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy    1021 CAACCACACACTTCCTAGAACACAGGCATTTAAATGAATTTGCGAGAAAGTTTAAGT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1042 CAACCACACACTTCCTAGAACACAGGCATTTAAATGAATTTGCGAGAAAGTTTATGT 1101

Qy    1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCCAATTAGAGAAAAATACTGGCGCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGCTCCAATTAGAGAAAAATTCGGCGCA 1161

Qy    1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTCATTAATGAAGATCTT 1221

Qy    1201 ATTAATGTATTAGAACAAATTGATAATGCTGTATTTTATTTATTGGGTGCAGGTGATATT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1222 ATTAATGTATTAGAACAAATTGATAATGCTGTGTTTTATTTATTGGGTGCAGGTGATATT 1281

Qy    1261 CAAAAATTACAAAAATGCATAATTTAGATAAATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1282 CAAAAATTACAAAAATGCATAATTTAGATAAATTTAGATAAAATTAGGCATGAAAAATGCGTTTTAA 1335
```

```
RESULT 9
AAV74703
ID   AAV74703 standard; DNA; 2424 BP.
XX
AC   AAV74703;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #392.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
```

RESULT 10
ABN92833
ID   ABN92833 standard; DNA; 1329 BP.
XX
AC   ABN92833;
XX
DT   24-JUL-2002 (first entry)
XX
DE   Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:2296.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy; gene; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-00134001.
XX
PR   14-AUG-1997;   97US-0055779P.
PR   08-NOV-1997;   97US-0064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA, Bush D;
XX
DR   WPI; 2002-381255/41.
DR   P-PSDB; ABP40288.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT   polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS   Disclosure; SEQ ID NO 2296; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences can
CC   also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life cycle
CC   or inhibit S. epidermidis infection. N.B. The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site
XX
SQ   Sequence 1329 BP; 482 A; 180 C; 247 G; 420 T; 0 U; 0 Other;

Query Match   70.0%; Score 922; DB 6; Length 1329;
Best Local Similarity   81.4%; Pred. No. 1.3e-171;
Matches 1069; Conservative   0; Mismatches 245; Indels   0; Gaps   0;

Qy   1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAAATC 75

Qy   61 ATGCATGATTTAGGACATGAGTTCAAGGATCCGGATATTGAGAACTACGTATTACAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   76 ATGCATGACCTCGGTCATGAGTACAAGGTTCAGACATAGAATCATACGTTTTTACAGAA 135

Qy   121 GTTGCTCTTAGAAATAAGGGATAAAAATATTACCATTGATGCTAATAAACATAAAAGAA 180
          ||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db   136 GTTGCATTAAGAAATAAAGGGATTAAAAATTTTACCTTTTGATGCAAAATAATTACAAAA 195

Qy   181 GATATCGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 240
          ||||| |||| ||| ||||||||||||||||||||||| |||||||||||||||||||
Db   196 GAAATGTTGTCATCCAAGGTAATGCATTCCTGATAATCATGAAGAAATTGTTAGGGCA 255

Qy   241 CATCAATTGAAATTAGATGTGTGTAAGTTATAATGATTTTTAGGACACAGATTATTGATCAA 300
          ||||||||||||||| |||| ||| |||||||||| |||||| |||||
Db   1887 CAAAAATTACAAAATGCATAATTAGATAAAATTAGATAAATTGAAAAATGCCGTTTTAA 1940

Db   807 GATATGGTAGTTATACAAGGTAATGCATTCGCCGGAGTAGCCATGAAGAAATAGTACGTGCA 866
Qy   241 CATCAATTGAAATTAGATGTGTGTAAGTTATGATCAA 300
          ||||||||||||||| |||| ||| |||||||| |||
Db   867 CATCAATTGAAATTAGATGTGTGTAAGTTATGATCAA 926

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   927 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 986

Qy   361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   987 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 1046

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1047 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 1106

Qy   481 TATAAACCTGATTACGCAATTATGACAAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||| ||||
Db   1107 TATAAACCTGATTACGCAATTATGACAAAATATTGNNNNNNNNNNNNNNNNNNNNNNNNNN 1166

Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
          ||||||||||||| |||||||||||||||||
Db   1167 NNNNNNNNNNNNNNNNNNNNNNNNNNNTGGCACATAATGTTAAAAAAGGTATT 1226

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1227 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 1286

Qy   661 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATTACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1287 TATGGATTTAAAGATTCGGATGACATTTATGTCTCAAAATATTCAAAATTACGGATAAAGGT 1346

Qy   721 ACTGCTTTTGATGTGTATGTGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 ACTGCTTTTGATGTGTATGTGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 1406

Qy   781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAAGAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAAGAGCTA 1466

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTAAACGTCGTTTCAAT 1526

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 1586

Qy   961 AGTGGCTACAAGTGAAACAGCACGAAAGAAAATTCCACATAAAGAAGTTGTTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 AGTGGCTACAAGTGAAACAGCACGAAAGAAAATCCACATAAAGAAGTTGTTGCAGTATTT 1646

Qy   1021 CAACCACACACTTCCTAGAACACAGGCATTTTAAATGAAATTGCAGAAAAGTTAAGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 CAACCACACACTTCCTAGAACACAGGCATTTTAAATGAAATTGCAGAAAAGTTAAGT 1706

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAAATTAGAGAAAATACTGGGCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAAATTAGAGAAAATACTGGCGCA 1766

Qy   1141 TTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1767 TTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT 1826

Qy   1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATGGTGTCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATGGGTGTCAGGTGATATT 1886

Qy   1261 CAAAAATTACAAAATGCATAAATTAGATAAAATTAGATAAAATGCCGTTTTAA 1314
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1887 CAAAAATTACAAAATGCATAATTAGATAAAATTAGATAAAATTGAAAAATGCCGTTTTAA 1940

ID   ACA47124 standard; DNA; 1311 BP.
XX
AC   ACA47124;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #28781.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU43254.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 34994; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1311 BP; 473 A; 181 C; 243 G; 414 T; 0 U; 0 Other;

```
Db       256  CATGAATTAAAGCTTGATATTATATAAAATATCATGACTTCTTGGTCATGTTATAAATCAA  315
Qy       301  TATACTTCAGTAGCTGTAACTGGTGCACATGTAAAACTTCTACACAGGTTTATTATCA     360
              ||||||| |||||||||||||||| ||||| |||||||||  ||| |||||
Db       316  TATACATCTGTTGCTGTCACAGGTGCACACGTGAAAACGTCTACAACGTGGTTTGTTGTCA  375

Qy       361  CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA    420
              |||||| |||||||| |||||| |||||||||| ||||||||||||||| ||||||||
Db       376  CATGTTGATGAATGGAGAGATAAAAAGACATCATTCCTAATGGTGATGGTACAGGTATGGGA  435

Qy       421  TTGCCTGAAAGTGATTATTCCGCTTTGAGGCATGTGAATAATAGACGTCACTTTTTAAGT  480
              |||||||       ||| ||||||  |||||||||||||||
Db       436  CTACCAGGAAGTGACTATTTGCTTTTGAGGCGTTGAATACCGTAGGCACTTCTTAAGT     495

Qy       481  TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
              |||||| |||||| |||||||| ||||||||||||||  |||||||||||||||
Db       496  TATCATCCAGATTATGCAATTATGACTAATATTGATTTTGACCATCCAGATTATTTTAAA  555

Qy       541  GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAGGTATT  600
              ||| ||| |||||||||||||||| ||||||  |||||||||| ||| ||||||||
Db       556  AATATCGATGATGTTTTATGATGCATTCCAACATCATAGCGTTAAATGCGTTAAGAAGGATT  615

Qy       601  ATTGCTTGGGGTGATGATGAACATCTCACGTAAAAATGAAGCAGATGTTCCAATTTATTAT  660
              ||||||||||||||||||| ||||  |   ||||||||| |||||||||||||||
Db       616  ATTGCTTGGGGTGATGATGTAATTTACGTAAATTAGATGTAGATATTCCGGTTTATTAT    675

Qy       661  TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT  720
              |||| |||||||||| |  ||| ||||| |||||||||  |||||||||
Db       676  TATGGCTTTAAAGAAACAGATGACATCTATGCTGAAGTAAAAATATTCAAATTACTGAAAACGGGT  735

Qy       721  ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCGTGTCTCCACAATAT  780
              ||| |||||||||| |||| |||| |||||   |||| ||  | ||||
Db       736  ACGCAATTTGATGTATATATTAAAGGCGAATTTATGATCAATTCTTATCCCCACACATAT  795

Qy       781  GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA  840
              ||||| |||| || |||| || ||||||| ||||| |||  || |   |||||||
Db       796  GGAAATCACAACATTCTAAACGCACTTGCTGTAATAGCAATTAGTTATTTAGAAAACATG  855

Qy       841  GATGTTACAAATATATAAAGAGAGCATTAAAGAAACGTTGGTGGTGTAAACGTCGTTCAAT  900
              ||||||   ||||||||||| |  |||| |||||||| || |||||||  | ||  |||
Db       856  AATGTAGAAAATATTAAAGAGCACTGATAACATTGGTGGTGTAAACGTCGTTTTAAT  915

Qy       901  GAAACTACAATTGCAAATCAAGTTAGTTGTAGATGAGTTATGCAGAGAAATT  960
              |||||||||| |||||| ||||  ||| |||| |||||||||| |||||||||
Db       916  GAAACAAAGTATCAAATCAAGTAGTAGTAGAGTAGTTATGCAGAGTATGCCAAGACAAGTTCAAGAGAGAAATT  975

Qy       961  AGTGCTACAATTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT  1020
              |||||||||||||||||  |||| |||| ||  | |||||||
Db       976  AGTGCTACAATTGAGAGACGCAGCGCCGAAAAAAAAAATATCCGCAAAAAGATGTTGTTGTCCGTTTTT  1035

Qy      1021  CAACCACACTTCTCTAGAACACAGGCATTTTAAAATGAATTTGCAGAAAGTTTAAGT  1080
              |||||||||||||||||  ||||||  |||||||||||| || ||
Db      1036  CAACCACATCGTTCTCAAGAACTCAAGCATTTTAAACGAATTTGCTGAGGTTTAAGT  1095

Qy      1081  AAAGCAGATCCTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAAATACTCGGCGCA  1140
              |||||||||||| ||| ||||||  | |||||||  ||||
Db      1096  AAGGCAGGACCAAGTATTTTTATGTGAAATATTCGGTTCAATAAGAGAAAATACGGGGAGAT  1155

Qy      1141  TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200
              ||||| | ||| | ||||| ||||| ||||||||| ||| | ||| |||||||
Db      1156  TTAACTATAGAAGATTAATCAATCGTATTGACGGCTTCGACGCCTTATAGATGAAAATAGT  1215

Qy      1201  ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTGATCGTCGTAGGTGGTCGAGGTGATATT  1260
              ||||||||||||||||||| |||||| |||| ||| |||||||||||||
Db      1216  ATTGATGTATTAGAAAAATTTGATAACGTCTGTAATTTTATTTTATTTATGGCGCGTCGGTGGTGGACATT  1275

Qy      1261  CAAAATTACAAAATGCATATTTAGATAATATTAGGCATGAAAAATGCGTTTAA  1314
              ||||| ||| || || || | ||| |||||||||||||||
Db      1276  CAAAAGTTACTTCAAAAGCTTATTTCGAAAAATTTGAAAAAAATTAGGCGCTAAAAAAATGATTTTTTAA  1329
```

Query Match   69.5%;   Score 915.8;   DB 7;   Length 1311;
Best Local Similarity   81.2%;   Pred. No. 2.2e-170;
Matches 1064;   Conservative   0;   Mismatches 247;   Indels   0;   Gaps   0;

```
Qy     1 ATGACACACATCATTTGTCCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC    60
         ||||||||||||||||| ||||||||||||||| |||||||| ||||| |||||||||||
Db     1 ATGACACACATCATTTGTCGGAATTAAAGGCTCAGGCATGAGTTCATTAGCACAAATC    60

Qy    61 ATGCATGATTTAGGACATGGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA   120
         ||||||||||||||| |||| ||||||||| ||||| |||||| |||||||||||| ||||
Db    61 ATGCATGACCTCGGTCATGAAGTTCAGACACAGGTTCAGACGATCATACGTTTTTACAGAA   120

Qy   121 GTTGCTCCTTAGAAATAAGGGATAAAATATTACCATTTGATGCTAATAACATAAAAGAA   180
         |||||||||||||||||| |||||||||||||  |||||||||||||||||||||||||
Db   121 GTTGCATTAAGAATAAAGGGATTAAAATTTTACCTTTTGATGCAAATAATATTACAAAA   180

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGGCAGTAGCCATGAAGAAAATAGTACGTGCA   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAAATGGTTGTCATCCAAGGTAATGCATTTCCTGATAATCATGAAGAAAATTGTTAGGGCA   240

Qy   241 CATCAATTGAAATTAGATGTTGTAAGTTATATAATGATTTTTAGGACAGATTATTGATCAA   300
         |||||||||||||||||||||| |||||| ||||||||||| |||||||| ||||||||||
Db   241 CATGAATTAAAGCTTGATATTTATTAAAAATCATGACTTTCTTGGTCATGTTATAAATCAA   300

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA   360
         |||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db   301 TATACATCTGTTGCTGTTGCTACAGGTGCACACGGTAAAACGTCACAACTGGTTTGTTGTCA   360

Qy   361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   420
         ||||||||| |||||||||||| ||||||||||||||| |||||| |||||||||||||
Db   361 CATGTGATGAATGGAGAGTAAAAGACATCATTTCCTAATCCGGTTACAGGTATGGGA   420

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAAGT   480
         |||||||||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db   421 CTACCAGGAAGTGACTATTTTGCTTTTGAGGCCTGTGAATACACCGTAGGCACTTCTTAGT   480

Qy   481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA   540
         ||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db   481 TATCATCCAGATTATGCAATTATGACTAATATTGATTTTGACCATCCAGATTATTTTAAA   540

Qy   541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT   600
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   541 AATATCGATGATGTTTTTGATGCATTCCAACATGCATTAAATGTTAAGAAGAAGGAATT   600

Qy   601 ATTGCTTGGGGTGATGATGAACATCACGCTAAAAATTGAAGCAGATGTTCCAATTTATTAT   660
         ||||||||| |||||||| |||||||||||||||||| ||||||||| ||||||||||||
Db   601 ATTGCTTGGGGGCGATGATGAATATTACGTAAATTAGTGGATATTTCCGGGTTTTATTAT   660

Qy   661 TATGGATTAAAGATTCGGATGACATTTTATGTCTCAAAATATTCAAATTACGGATAAAGGT   720
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db   661 TATGGCTTTAAAGAAACAGATGACATCTATGCTAAAAATATTACTCAAATTACTGAAAAGGGT   720

Qy   721 ACTGCTTTTGATGTGTATGTCGATGGTGAGTTTATGATCACTTCCTGTCTCTCCACAATAT   780
         ||||||||||||||||  ||||||||||||||||||||||||||| ||||||||||||||
Db   721 ACGCAATTGATGTATATATTAAAGGCGAATTTATGATCAATTCTTATCAATTCTTATCCCCAATAT   780

Qy   781 GGTGACCATACAGTTTAAATGCATTTAGCTGTAATTGCGATTAGTAGTTATATTTAGAAGCTA   840
         |||||||| ||||||||||||| ||| ||||||| |||||||||||||||||| ||||||
Db   781 GGAAATCACACACATTCTAAACGCACTTGCTGTAATAGCAATTAGTATTAGAAGACATG   840

Qy   841 GATGTTACAAATATTAAAGAAGCACTGAATAACATTTGGTGGTTAAAACGTCGTTTCAAT   900
         ||||||||||||||||||||||||||||||| |||||||||||||||| |||||||| |||
Db   841 AATGTAGAAAATATTAAAGAAGACACTGATAAACAATTTGGTGGTGTAAAACGTCGTTTTAAT   900

Qy   901 GAAACTACAATTGCAAATCGAATCAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT   960
         |||| |||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db   901 GAAACAAAGTATCAAATCAAATCAAGTATCAAATAGTAGAGACCGATTATGCACCCATCCAAGAGAAATT   960

Qy   961 AGTGCTACAATTGAACAGCACGAAAGAAAATATCCATAAAGAAGTTGTTCAGTATTT   1020
         |||||||||||||||| ||||| |||| |||||||||||| |||||||||||||||||
Db   961 AGTGTCTACAATTGAGACAGCAGCGCGAAAAAAAATACCGCAAAAAGATGTTGTTGCCGTTTTT   1020

Qy  1021 CAACCACACACTTCTCTAGAACACAGGCATTTAAATGAATTGCAGAAAGTTAAGT   1080
         |||||||||||||||||| |||||||||||||||  |||||||| ||||||||||
Db  1021 CAACCATACACGTTCTCAAGAACTCAAGACTCAGCATTTTAAACGAATCAAGCATTTGCTGAGAGTTTAAGT   1080

Qy  1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA   1140
         || |||||||||| ||||||||||||| ||||||||||| |||||| ||||||||||||
Db  1081 AAGGCAGCAGCCAAGTATTTTTATGTGAAATATTCGGTTCAATATTCGGTTCAATAATATTACGGAGAGAT   1140

Qy  1141 TTAACGATACAAGATTAATGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  1141 TTAACTATAGAGATTAATCAATCATTGATTACGGCTCGACGCTTATAGATGAAAATAGT   1200

Qy  1201 ATTAATGTATTAGAACAATTTGATAAATGCTGTTATTTTATTTATGGGGTGCAGGTGATATT   1260
         |||  |||||||| |||||||||||||| ||||||||||| |||||||| |||||||||||
Db  1201 ATTGATGTATTAGAAGAAAATTTGATAACGCTGTAATTTTATTTATGGGGGCCTGGTGACATT   1260

Qy  1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGTGAAAAATGCGTTT   1311
         ||| ||||| |||||||||  |||||||||||||||||||||||||||
Db  1261 CAAAAGTTACTAAAAGCTTATTTCGAAAATATTAGGGCGTAAAAAATGATTTT   1311
```

RESULT 12
AAV80065
ID   AAV80065 standard; DNA; 660 BP.
XX
AC   AAV80065;
XX
DT   17-MAR-1999 (first entry)
XX
DE   Partial nucleotide sequence of the MurC gene.
XX
KW   MurC gene; UDP-N-acetylmuramate:L-alanine ligase; MurC polypeptide;
KW   bacterial; infection; H. pylori; cancer; ulcer; gastritis; vaccine;
KW   immunogen; drug; genetic immunisation; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key   Location/Qualifiers
FT   CDS   2..660
FT   /tag= a
FT   /product= "partial MurC polypeptide"
XX
PN   EP889123-A2.
XX
PD   07-JAN-1999.
XX
PF   26-JUN-1998;   98EP-00305064.
XX
PR   03-JUL-1997;   97US-0052720P.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
PA   (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI   Wallis NG, Burnham MKR;
XX
DR   WPI; 1999-062655/06.
DR   P-PSDB; AAW89199.
XX
PT   New isolated MurC polypeptide from Staphylococcus aureus and related
PT   nucleic acid - useful in diagnosis, treatment and prevention of bacterial
PT   infections.
XX
PS   Claim 2; Page 4-5; 39pp; English.
XX
CC   The invention relates to a UDP-N-acetylmuramate:L-alanine ligase (MurC
CC   polypeptide) encoded by the S. aureus MurC gene. Host cells containing an
CC   expression system comprising the MurC gene can be used for the
CC   recombinant production of the polypeptide. Agonists or the MurC
CC   polypeptide are used to treat conditions requiring increased activity or
CC   expression of the polypeptide. Antagonists, inhibitory nucleic acid or
CC   competitive polypeptide are useful for inhibiting the polypeptide e.g.
CC   bacterial (especially S. aureus) infections. They are also useful against
CC   Helicobacter pylori infections and related cancers, ulcers and gastritis.
CC   The antibacterial agents are useful to treat in-dwelling devices for

```
CC   infection prevention or generally as wound treatments to prevent adhesion
CC   of bacteria to matrix proteins. The MurC polypeptide is also useful for
CC   diagnosing or prognosing a (susceptibility to) disease, for raising
CC   antibodies; to identify modulators or specific receptors; in rational
CC   drug design and as an immunogen for vaccines. The MurC gene sequences are
CC   useful in antisense/ribozyme therapeutics; to detect mutant MurC gene;
CC   for chromosomal mapping; to determine bacterial serotype; and for genetic
CC   immunisation. The present sequence represents a partial nucleotide
CC   sequence (MurC ORF) of the MurC gene
XX
SQ   Sequence 660 BP; 233 A; 83 C; 120 G; 224 T; 0 U; 0 Other;

Query Match             48.0%;  Score 633;  DB 2;  Length 660;
Best Local Similarity   98.5%;  Pred. No. 6.1e-115;
Matches  639;  Conservative   0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    666 ATTTAAAGATTCGGATGACATTATGTCTCAAAATATTCAAAATTACGGATAAAGGTACTGC 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTTAAAGATTCGGATGACATTATGTCTCAAAATATTCAAAATTACGGATAAAGGTACTGC 60

Qy    726 TTTTGATGTGTATGTGGATGTGAGTTTATGATCACTTCCTGTCTCCACATATGTGTGA 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTTGATGTGTATGTGGATGTGAGTTTATGATCACTTCCTGTCTCCACATATGTGTGA 120

Qy    786 CCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTAGATGT 845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTAGATGT 180

Qy    846 TACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTCAATGAAAC 905
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    181 TACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTCAATGAAAC 240

Qy    906 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 965
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGC 300

Qy    966 TACAATTGAAACAGCACGAGAAGAAATATCCACATAAAGAAGTGTTGCAGTATTTCAACC 1025
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TACAATTGACACAGCACGAGAAGAAATATCCACATAAAGAAGTGTTGCAGTATTTCAACC 360

Qy   1026 ACACACTTTCTCTAGAACAACAGGCATTTTAAATGAATTTTGCAGAAAGTTAAGTAAAGC 1085
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    361 ACACACTTTCTCTAGAACAACAAGCATTTTAAATGAATTTTGCAGAAAGTTTATGTAAAGC 420

Qy   1086 AGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCATTAAC 1145
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    421 AGATCGTGTATTCTTATGTGAAATTTTTGGCTCAATTAGAGAAAATTCTGGCGCATTAAC 480

Qy   1146 GATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATCTTATTAA 1205
          |||||||||||||||||||||||||||||||||||||||||||| |||||
Db    481 GATACAAGATTAATTGATAAAATTGGAGGTGCATCGTTAATTAATGAAGATCTTATTAA 540

Qy   1206 TGTATTAGAACAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATATTCAAAA 1265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TGTATTAGAACAATTTGATAATGCTGTTGTTTTATTTATGGGTGCAGGTGATATTCAAAA 600

Qy   1266 ATTACAAAATGCATATTAGGACATGAAAAATGCGTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||
Db    601 ATTACAAAATGCATATTAGGACATGAAAAATGCGTTTAA 649
```

```
RESULT 13
ACA22277
ID   ACA22277 standard; DNA; 1308 BP.
XX
AC   ACA22277;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #3934.
XX
KW   Antisense; db; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
```

```
XX
OS   Bacillus anthracis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002;  2002WO-US009107.
XX
PR   21-MAR-2001;  2001US-00815242.
PR   06-SEP-2001;  2001US-00948993.
PR   25-OCT-2001;  2001US-0342923P.
PR   08-FEB-2002;  2002US-00072851.
PR   06-MAR-2002;  2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,     Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,     Trawick JD,  Carr GJ,    Yamamoto R,    Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU18407.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 10147; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1308 BP; 456 A; 231 C; 267 G; 354 T; 0 U; 0 Other;

Query Match             45.3%;  Score 597.6;  DB 7;  Length 1308;
Best Local Similarity   66.4%;  Pred. No. 5.9e-108;
Matches  858;  Conservative   0;  Mismatches 434;  Indels   0;  Gaps   0;

Qy      1 ATGACACACTATCATTTGTCGGATTAAAGGTTCTGGCATGAGTTCATTAGAGTTCATTAGCACACAAATC 60
          | |||||| | ||||||  ||| |  |  ||  |  ||||   || |||| | ||||||||||
Db      1 ATGACAGTTTACCATTTGTGAGGAATTAAAGGAACAGGAATGAGTTCATTAGCGTATTACGGCAAATT 60

Qy     61 ATGACATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACCAGAA 120
          | |||||| | ||   | || | |||  |||   ||||||| |||  ||||| || |||||  ||
Db     61 ATGACGTGTTTATTTGTGTTGTTTATTTGGGTGCAGGTGCGCAGTATTGAGAACTACGTATTTACCAGAA 120
```

```
Db   61 .............................................................
Qy   61 CTTCATGACATGAAGCATACTGTTCAAGGGTCTGATTATGAAAAGGCGTTTCTTACACAA 120

Qy  121 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
Db  121 ACAGCCGTTGGAAAAGCGTAATAATCTCGATTCTTCCTTTGATAAAAGTAATGTAAAAGAA 180

Qy  181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAAAATAGTACGTGCA 240
Db  181 GGACAAGTGATTATTGCAGGAAATGCATGCATTCCTGATACGCGATGAGAAATCGTAGCAGCA 240

Qy  241 CATCCAATTGAAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATATTGATCAA 300
Db  241 AAAGAATTAAACATCCCAGTACATCGTTACCATCACTTCCTTAGGTGATCTTATGAACCAA 300

Qy  301 TATACTTCAGTAGCTGGTGCACATGGTAAAACTTCTACACAGGTTTATTATCA 360
Db  301 TACACAAGTCGTTGCTGTAACTGGTGCGCGCATGGCTGATACAACAACTGGTTTGTTGTAGCC 360

Qy  361 CATGTTATGAAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
Db  361 CATGTAAATCGAAGGTGCACACCCTACATCTTACCTTATTGGAGATGGAACAGGGCATGGG 420

Qy  421 TTGCCTGAAAGTGATTATTCGCCTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
Db  421 GTAGAAAATAGTAAGTATTTTGTATTTGTAGGCTTGTGAGTATCGTCGTCATTTCTTGTCT 480

Qy  481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
Db  481 TACAATCCAGACTACTTGCAATTATGACAAACATTGATTTTGATCATCCAGATTATTCACA 540

Qy  541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
Db  541 GATATCAATGATGTATTCAGTGCATTCCAAGAGATGGGCATTGCAAGTGAAAAAAGGCATT 600

Qy  601 AATTGCTTGGGGTGATGATGAACATCACGTAAGCAGATGTTCCAATTTATTAT 660
Db  601 AATTGCATGTGGAGATGATGAAGAACTTCAAAGAACTTCAAGCGAAAGTACCTGTTATTTTC 660

Qy  661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGGT 720
Db  661 TATGGATTTGGGAGAAGATATATGATTTTCCAAGCACGTAACATTCAAAAGAGAACTGACGGT 720

Qy  721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCGTCTCCACAATAT 780
Db  721 ACTATTTTCGATGTTTCGTTCGTAATACGTACTATGACACGTTCAAAATTACAGGCTAC 780

Qy  781 GGTGACCATACAGTTTAAATGTTTAAAATGCATTAGTCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 840
Db  781 GGCAACCACAGCGTATTAAATGCATTAGCAGTAATCGCACTTTGCCATTATGAAAACGTT 840

Qy  841 GATGTTACAAAATATTAAAGAAGCAGCATTAGAAACGTTTGGTGGTGTAACGTCGTTTCAAT 900
Db  841 GATGTAGAAGCAGTTAAGCATCAGTTAACAACTTTGAAGGCGTAAAACGTCGCTTTAAT 900

Qy  901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAGAAATT 960
Db  901 GAAAAGCCAATGGGAGAGCAAGTTATTATTTATGATGACTACGGCACCATCCGACAGAAATT 960

Qy  961 AGTGCTACAATTGAAACAGCACGAAAGAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
Db  961 AATGCAACGATTGAAGCAGCTCGTCAAAAAACATCCAGACGTGAAATTGTCGCTGTATTC 1020

Qy 1021 CAACCACCACTTCTCTAGAACAACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
Db 1021 CAGCCGCACACAATTCTCACGTACAGAAAAGTTCTTTAGATGAGTTCGCTGAAAGCCTAAGC 1080

Qy 1081 AAAGCAGATCGTGATTCTTATGTGAAAATTTTTGAAGATTTTTTCCGATCGAACGTTGGCGCA 1140
Db 1081 AAAGCTGACCAAGTATACTTATGTGATATTTTCGGATCCAGCCGGATCAGGAAAACAAAGTGAA 1140

Qy 1141 TTAACGATTACAAGAATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATTGAAGATTCT 1200
Db 1141 TTAACAATCGAAGGATCTCGCAAAAGCGTATTGACGGTCGCAGAACTAATTACAGATACAACA 1200
```

```
Qy 1201 ATTAATGTATTAGAACAATTGATAATGCTGTATTTTATTTATGGGTGCAGGTGATATT 1260
Db 1201 ACGGATGTATTTAAAGAAACATAAAAACGGCGTTCCATTTCATGGCGGCCAGGAGACATC 1260

Qy 1261 CAAAAATTACAAAATGCATATTTAGATAAATT 1292
Db 1261 CAAAAATTCGAAGCAGCTTACGTAACGAAAGAAGT 1292
```

RESULT 14

```
AAV53479/c
ID  AAV53479 standard; DNA; 619 BP.
XX
AC  AAV53479;
XX
DT  30-OCT-1998 (first entry)
XX
DE  DNA encoding a Staphylococcus aureus protein of unknown function.
XX
KW  Staphylococcus aureus protein; immune response induction; eye infection;
KW  antibody production; T-cell immune response; gastrointestinal infection;
KW  respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW  central nervous system; kidney infection; urinary tract infection;
KW  antimicrobial compound identification; broad spectrum antibiotic;
KW  therapy; ss.
XX
OS  Staphylococcus aureus.
XX
PN  EP841394-A2.
XX
PD  13-MAY-1998.
XX
PF  24-SEP-1997;  97EP-00307485.
XX
PR  24-SEP-1996;  96US-0027032P.
XX
PA  (SMIK ) SMITHKLINE BEECHAM CORP.
PA  (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI  Black MT, Hodgson JE, Knowles DJC, Reichard RW, Nicholas RO;
PI  Burnham MKR, Pratt JM, Rosenberg M, Ward JM, Lonetto MA;
XX
DR  WPI; 1998-252940/23.
DR  P-PSDB; AAW77686.
XX
PT  New nucleic acid sequences from Staphylococcus aureus WCHU29 - useful in
PT  vaccines and for treatment of bacterial infections of e.g. respiratory
PT  tract and central nervous system.
XX
PS  Claim 1; Page 146; 390pp; English.
XX
CC  This sequence encodes a Staphylococcus aureus protein of unknown
CC  function, and represents a DNA Sequence of the invention. The DNA
CC  sequences were isolated from Staphylococcus aureus WCHU29 (NCIMB 40771).
CC  Host cells containing the DNA sequences are used to produce polypeptides
CC  or fragments. The proteins are used in the treatment of disease, for
CC  inducing an immune response by administering them, to produce antibody
CC  and/or T-cell immune response. Antagonists of the proteins are used for
CC  the inhibition of bacterial polypeptides. Conditions which may be treated
CC  include bacterial infection, especially respiratory, cardiac,
CC  gastrointestinal, central nervous, eye, kidney, urinary tract, skin,
CC  bones and joints. The proteins can also be used to identify antimicrobial
CC  compounds which are broad spectrum antibiotics, especially useful in the
CC  treatment of H. pylori infection
XX
SQ  Sequence 619 BP; 208 A; 117 C; 79 G; 215 T; 0 U; 0 Other;

Query Match             43.3%;  Score 570.4;  DB 2;  Length 619;
Best Local Similarity   98.0%;  Pred. No. 1.2e-102;
Matches  588;  Conservative    0;  Mismatches   11;  Indels    1;  Gaps    1;

Qy  716 AAGGTACTGCTTTTGATGTGATGTGGATGGTGAGTTTTATGATGACTTCCTGTCTCCAC 775
```

```
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU35277.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 6773; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1371 BP; 436 A; 236 C; 296 G; 403 T; 0 U; 0 Other;

Query Match             42.1%;  Score 555;  DB 4;  Length 1371;
Best Local Similarity   64.7%;  Pred. No. 1.4e-99;
Matches   841;  Conservative   0;  Mismatches  455;  Indels   3;  Gaps   1;

Qy    10  TATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATCATGCATGAT   69
          ||||||||||||||||||||||||||||||||||||||||||||
Db    58  TACCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCGTTGGTCTTACACCAA  117

Qy    70  TTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACAACACGTATTACAGAAGTTGCTCTT  129
          |||||||   |||||||||||||| ||| ||||| ||       ||||||  |  | ||
Db   118  AAAGGCTACAACGTTCAAGGATCAGATGTAGAAGAGTATTTCTTTACAACAACGGGGATCTT  177

Qy   130  AGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAAGAGATATGGTA  189
          | |||| |||| | || ||| || |||||  |||  |||| |  | ||||| |||||| ||
Db   178  GAAAAATCAGGTGTCCCTATTTACCATTTAACCATTTAATGCAGATAATATTGATAAAGACATGATT  237

Qy   190  GTTATACAAGGTAAGTCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCAATCATTG  249
          |||||  | || ||  |||   |    || ||||  |||||  |     ||  |  || 
Db   238  GTTATTGCTGGAAATGCTTTTCCAGATACCCATGGAAATCGCCCGGCCATCGGAATTA  297

Qy   250  AAATTAGATGCTGTGAAGTTATAATGATTTTTTAGGACAGATTGATCAATATACTTCA  309
          |  |  || | || | | ||| |||| | |||  || ||| |||  ||| ||| | ||
Db   298  GGCGCAGAGAGTAAATTCGTTACCACGATTCATTGCTCGTTTATCGAACCGTACACAAGC  357

Qy   310  GTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACACAGGTTTATTATCACATGTTATG  369
          ||||||||||| ||  || ||||| ||||   | || |   |||||| || ||||||| | |
Db   358  ATTGCTGTAACAGGGTCACATGGCGAAAACAAGTACGACTCGTTGTTGCTAGCACATGTATTA  417

Qy   370  AATGGTGATAAAAAGACTTCATTTTAATGGTGATGGCACAGGTATGGGATTGCCTGAA  429
          | ||||||| |   |  ||  ||||| | | |||||||| || |  || |  || || 
Db   418  AGTGTATCAATCCAACTAGTTATTTAATAGGGGATGGCACTGGCCACGGCGAACCAGAT  477

Qy   430  AGTGATTGATTATTCGCTTTTGAGGCATGTGAATGTGAATATGACGTGCACTTTTTAAGTTATAAACCT  489
          |  ||| | ||| |  |||| |||||| || || || |||| || |  |||   || ||||||||||||
Db   478  GCTGATTTCTTTGCATTTGAAGCGTGTGAATATCGCCGTCATTCTGGCTTATTCACCA  537

Qy   490  GATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTAAAGATATTAAT  549
          |||||||| |||||||||||||| ||| | || || |||||| ||||| |||| | ||
Db   538  GATTATGCGATTATGACGAATATCGATTTTGATCATCCAGATTACTACAAGCATTGAG  597

Qy   550  GATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGTATTATTGCTTGG  609
          |||||||| || ||||||||| ||||||| || || || |||||| |||||||| || 
Db   598  GACGTTTTTTCAGCGTTCCAAACATGGCTCATCAAGTCAAAAAGGAATTTTGCTTAT  657
```

```
Db   616  |  |||||||||||||||||||||||||||||||||||||||||||||  ACGAGACTGCTTTGATGTGTATGTGTAAAATGCATTAGCTGTAATTGCCGATTAGTATTTAGAGA  557
Qy   776  AATATGGTGACCATACAGTTTAAAATGTGTATGTGTAAAATGCATTAGCTGTAATTGCCGATTAGTATTTAGAGA  835
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   556  AATATGGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCCGATTAGTATTTAGAGA  497
Qy   836  AGCTAGATGTTACAAAATATTAAAGAAGACATTAGAAACGTTTGGTGGTGTTAAACGTCGTT  895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496  AGCTAGATGTTACAAAATATTAAAGAAGACATTAGAAACGTTTGGTGGTGTTAAACGTCGTT  437
Qy   896  TCAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAG  955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   436  TCAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAG  377
Qy   956  AAATTAGTGCTACAAATGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAG  1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376  AAATTAGTGCTACAAATTGACACAGCACGACAGAAATATCCACATAAAGAAGTCGTTGCAG  317
Qy  1016  TATTTCAACCACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTT  1075
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   316  TATTTCAACCACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTT  257
Qy  1076  TAAGTAAAGCAGCGATGCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTG  1135
          ||||||||||||||||||||||||| |||||||||||||||||||| |||||||||||||
Db   256  TATGTAAAGCAGATCGTGTATTCTTATCTTATGTGAAATTTTGGCTCAATTAGAGAAAATTCTG  197
Qy  1136  GCGCATTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAG  1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   196  GCGCATTAACGATACAAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAG  137
Qy  1196  ATTC-TATTAATGTAATTAGAACAATTGATATATGCTGTTATTTTATTGGTGTGCAGGT  1254
          |||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   136  ATTCCTATTAATGTAATTAGAACAATTGATATATGCTGTTATTTTATTGGTGTGCAGGT  77
Qy  1255  GATATTCAAAAATTACAAAAATGCATATTTTAGATAAATGAAAAATGCGGTTTTAA  1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    76  GATATTCAAAAATTACAAAAATGCATATTTTAGATAAATGAAAAATGCGGTTTTAA  17
```

```
RESULT 15
AAS53136
ID   AAS53136 standard; DNA; 1371 BP.
XX
AC   AAS53136;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis DNA for cellular proliferation protein #564.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US0009180.
XX
PR   21-MAR-2001; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
```

```
Qy    610 GGTGATGATGAACATCTACGTAAAATTGAAGCAGAGATGTTCCAATTTATTATTATTATGGATTT 669
          ||||||||||| |||||| ||||   ||||||||   ||||||||||  ||||||||||||||
Db    658 GGTGATGATAAGTATCTTCGCCAGTTAGAATCAGAAGTGCCAGTTTATTTATTTATTATGGGGTC 717

Qy    670 AAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGTACTGCTTTT 729
          |  |||||| ||| |||  ||| |||||||| ||  |||||||||||| ||||  |||||
Db    718 AGCGAAGAGGATGATATCCAAGCCGAAATATTCAACGAACAACGGAAGGCTCATCTTTT 777

Qy    730 GATGTGTATGTCGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGTGACCAT 789
          ||||||| ||| ||||  ||||| |||||   ||||||| ||||| ||||| ||||||
Db    778 GATGTTTATCACAAGGCTGATTTTGTAGGTCATTTTGTCTTACCAGCATTTGGCCATCAC 837

Qy    790 ACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTGATTAATTAGAGAAGCTAGATGTTACA 849
          ||||||   ||||| |||| |||| |||  |||| ||||  ||||| ||||||||||  ||||
Db    838 AATATCATGAATGCGCAGGTGTGATTGCTGTGTGGCTTATTTTGAAAAACTTGATATGCAA 897

Qy    850 AATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAATGAAACTACA 909
          || ||||||||||  |||  ||||||  ||||||  ||||||||| |  |||||| ||||
Db    898 AAAGTCCGAGAAGAAATGCTAAGTTTTAAAGGTGTGAAACGTCGTTTTAGCCGAGAAAAAA 957

Qy    910 ATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGCTACA 969
          || ||||| |||  || ||| ||||||||| || |||| ||||| ||||||||||||||
Db    958 GTCAGTGACATGATTATTGTTGATGATTATGCGCACCATCCAGCTGAAATTAAAGCAACG 1017

Qy    970 ATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTTCAACCACAC 1029
          ||||||| ||| || ||| |||||||||||||| |||||| |||||| || |||| ||||
Db   1018 ATTGATGGCGCCCCGCCAAAAATATCCTGACAAAGAAATTATTGCTGTCTTCCAGCCACAT 1077

Qy   1030 ACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAGTTTAAGTAAAGCAGAT 1089
          || ||| |||||| || |||| |||  || ||||||||||| |||  ||||||||||
Db   1078 ACATTTACACGAACAATTGCCCTAAATGGATGAATTTGCTGAGCACTGGATTTGGCAGAT 1137

Qy   1090 CGTGTATTCCTATGTGAAAATTTTGGATCAATTAGAGAAAATACTGGCGCCATTAACGATA 1149
          |||| | |||  | || ||||  |||||| |||| ||||||| ||||| ||||| |  ||
Db   1138 GAAGTATTCTTATGTAATATATTTTGGCTCTGCGGTGAAACAAGGCGAGGTACGCATT 1197

Qy   1150 CAAGAATTTAATTGATAAAATT---GAAGGTGCATCGTTAATTAAATGAAGATTCTATTAAT 1206
          ||||||| |||| |||||||  | ||||  ||   |||| ||||| |||||| ||| |||
Db   1198 GAAGATTTAGGTGAAAAAATTCAAAAAAGGTCGACAAGTAATTACCGAAGATAATGTGTCG 1257

Qy   1207 GTATTAGAACAATTGATAATGCTGTTATTTTATTTTATGGGTGCAGTGATATTCAAAAA 1266
          || |||| ||| ||||| |||| |  |||  ||||| ||||| ||| || ||||||||
Db   1258 CCTTTACTAGATTTTGAAAATGCAGAGGTGGTCTTTTATGGGTGCTGGCGACGTTCAGAAA 1317

Qy   1267 TTACAAAATGCATATTTAGATAAAATTAGGCATGAAAAAT 1305
          |||| |||  |  |||  |||   ||||| |||  | |||
Db   1318 TTTGAACAAGCCTACGAAACATTACTAAGTAACAACACACT 1356
```

```
OM nucleic - nucleic search, using sw model
              GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:       June  5, 2004, 23:46:19 ; Search time 134 Seconds
                                     (without alignments)
                           5458.398 Million cell updates/sec

Title:        US-10-712-713-1
Perfect score: 1318
Sequence:     1 atgcacactatcatttgt.........gaaaatgcgtttaagctt 1318
Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0
Searched:     682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:     1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents NA:*
              1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
              6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES
```

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1314	99.7	2424	4	US-08-956-171E-392	Sequence 392, App
	2	922	70.0	1329	4	US-09-134-001C-2296	Sequence 2296, Ap
	3	570.4	43.3	619	4	US-09-936-165A-179	Sequence 179, App
	4	400.8	30.4	11864	4	US-08-961-527-61	Sequence 61, Appl
	5	399.2	30.3	2715	4	US-08-940-572-1	Sequence 1, Appl
	6	389.6	29.6	888	3	US-08-714-918-61	Sequence 61, Appl
	7	389.6	29.6	888	3	US-09-265-315-61	Sequence 61, Appl
c	8	389.6	29.6	888	3	US-09-265-315-61	Sequence 61, Appl
c	9	389.6	29.6	888	3	US-09-266-417-61	Sequence 61, Appl
c	10	389.6	29.6	888	4	US-09-528-709-61	Sequence 61, Appl
c	11	389.6	29.6	888	4	US-09-527-745-61	Sequence 61, Appl
	12	370.2	28.1	1267	3	US-08-961-083-115	Sequence 115, App
	13	370.2	28.1	1267	4	US-09-536-784-115	Sequence 115, App
	14	354.4	26.9	873	4	US-09-107-532A-3260	Sequence 3260, Ap
c	15	310.4	23.6	1825	4	US-08-940-572-3	Sequence 3, Appli
	16	138.6	10.5	1479	4	US-09-328-352-897	Sequence 897, App
c	17	105.4	8.0	640681	4	US-09-790-988-1	Sequence 1, Appli
	18	84.4	6.4	1467	4	US-09-543-681A-2779	Sequence 2779, Ap
	19	78.4	5.9	1830121	4	US-09-557-884-1	Sequence 1, Appli
	20	78.4	5.9	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	21	76.2	5.8	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	22	72.8	5.5	5361	3	US-08-973-462-2	Sequence 2, Appli
	23	72.8	5.5	6152	3	US-08-973-462-1	Sequence 1, Appli
	24	58.8	4.5	1485	3	US-09-543-681A-601	Sequence 601, App
	25	55.2	4.2	8920	2	US-08-446-855A-1	Sequence 1, Appli
	26	55.2	4.2	8920	3	US-09-150-741-1	Sequence 1, Appli
c	27	54.4	4.1	1830121	4	US-09-557-884-1	Sequence 1, Appli
c	28	54.4	4.1	1830121	4	US-09-643-990A-1	Sequence 1, Appli
c	29	54	4.1	15016	1	US-09-601-198-60	Sequence 60, Appl
	30	53.8	4.1	2223	3	US-08-257-073-4	Sequence 4, Appli
	31	53.6	4.1	1891	3	US-08-973-462-3	Sequence 3, Appli
	32	53.4	4.1	1482	4	US-09-540-236-792	Sequence 792, App
	33	53.4	4.0	19619	4	US-09-596-002-14	Sequence 14, Appl
c	34	52.8	4.0	5340	4	US-09-627-122-21	Sequence 21, Appl
c	35	52.4	4.0	14066	4	US-09-601-198-56	Sequence 56, Appl
	36	50.8	3.9	1374	4	US-09-134-001C-1447	Sequence 1447, Ap
	37	50.2	3.8	1416	4	US-09-543-681A-1533	Sequence 1533, Ap
	38	50	3.8	9636	1	US-08-323-170B-1	Sequence 1, Appli
	39	50	3.8	9636	1	US-08-954-441-1	Sequence 1, Appli
	40	49.8	3.8	2606	2	US-08-568-459A-7	Sequence 7, Appli
	41	49.8	3.8	2606	2	US-08-487-826B-7	Sequence 7, Appli
	42	49.8	3.8	2606	4	US-09-210-288-7	Sequence 7, Appli
	43	49.6	3.8	520	4	US-08-961-527-344	Sequence 344, App
	44	49.4	3.7	1692	4	US-09-601-198-63	Sequence 63, Appl
c	45	49	3.7	5652	4	US-09-601-198-75	Sequence 75, Appl

ALIGNMENTS

```
RESULT 1
US-08-956-171E-392
; Sequence 392, Application US/08956171E
; Patent No. 6559114
; GENERAL INFORMATION:
;  APPLICANT: Charles Kunsch
;             Gil H. Choi
;             Patrick S. Dillon
;             Craig A. Rosen
;             Steven C. Barash
;             Michael R. Fannon
;  TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;  NUMBER OF SEQUENCES: 5256
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/956,171E
;   FILING DATE: 20-Oct-1997
;   CLASSIFICATION: <Unknown>
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/009,861
;   FILING DATE: January 5, 1996
;   APPLICATION NUMBER: 08/781,986
;   FILING DATE: January 3, 1997
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Mark J. Hyman
;   REGISTRATION NUMBER: 46,789
;   REFERENCE/DOCKET NUMBER: PB248P1
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (240) 314-1224
;   TELEFAX: (301) 309-8439
;  INFORMATION FOR SEQ ID NO: 392:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH: 2424 base pairs
;    TYPE: nucleic acid
;    STRANDEDNESS: double
;    TOPOLOGY: linear
;  SEQUENCE DESCRIPTION: SEQ ID NO 392:
US-08-956-171E-392
```

Query Match   99.7%; Score 1314; DB 4; Length 2424;
Best Local Similarity 100.0%; Pred. No. 6.4e-296; Indels 0; Gaps 0;
Matches 1314; Conservative 0; Mismatches 0;

```
Qy      1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 ATGACACACTATCATTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC    686

Qy     61 ATGCATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 ATGCATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA    746

Qy    121 GTTGCTCCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAGAA    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 GTTGCTCCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAGAA    806

Qy    181 GATATGGTAGTTATACAAGTAATGTCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCA    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 GATATGGTAGTTATACAAGTAATGTCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCA    866

Qy    241 CATCAATTGAAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 CATCAATTGAAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA    926

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA    986

Qy    361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA   1046

Qy    421 TTGCCCTGAAAGTGATTATTTCGCTTTTGGAGGCATGTGAATATAGAGCGTCACTTTTTAAA    480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1047 TTGCCCTGAAAGTGATTATTTCGCTTTTGGAGGCATGTGAATATAGAGCGTCACTTTTTAAGT   1106

Qy    481 TATAAACCTGATTACGCCAATTATGATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1107 TATAAACCTGATTACGCCAATTATGATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA   1166

Qy    541 GATATTAATGATGTTTTTGATGCATTCCAAGAAAATGGCACATAATGTTAAAAAAGGTATT    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1167 GATATTAATGATGTTTTTGATGCATTCCAAGAAAATGGCACATAATGTTAAAAAAGGTATT   1226

Qy    601 ATTGCTTGGGGTGATGAGTGAACATCTCACGTAAAAATTGAAGCAGATGTTCCAATTATTAT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1227 ATTGCTTGGGGTGATGAGTGAACATCTCACGTAAAAATTGAAGCAGATGTTCCAATTATTAT   1286

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGTCAAAATATTCAAATTACGGATAAAGGT    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1287 TATGGATTTAAAGATTCGGATGACATTTATGTCAAAATATTCAAATTACGGATAAAGGT   1346

Qy    721 ACTGCTTTTGATGTGTATGTGGAGTTTATGATGAGTTTATTGCGATTAGTTATTTAGAGAAGCTA    780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 ACTGCTTTTGATGTGTATGTGGAGTTTATGATGAGTTTATTGCGATTAGTTATTTAGAGAAGCTA   1406

Qy    781 GGTGACCATACAGTTTTAAATGCATTAGCTCGATTAGTTATTTAGAGAAGCTA    840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 GGTGACCATACAGTTTTAAATGCATTAGCTCGATTAGTTATTTAGAGAAGCTA   1466

Qy    841 GATGTTACAAATATTAAAGAAGAGCATTAGAAACGTTTGGTCGTGTTAAACGTCGTTTCAAT    900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 GATGTTACAAATATTAAAGAAGAGCATTAGAAACGTTTGGTCGTGTTAAACGTCGTTTCAAT    926

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCCACATCCAAGAGAAATT    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCCACATCCAAGAGAAATT   1586

Qy    961 AGTGCTACAATTGAAACAGCACGAGAAGAAATATCCACATAAAGAAGAGTTGTTGCAGTATTT   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 AGTGCTACAATTGAAACAGCACGAGAAGAAATATCCACATAAAGAAGAGTTGTTGCAGTATTT   1646

Qy   1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAGAAAATACTGGCGCA   1706

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGACAGTTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCA   1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 AAAGCAGATCGTGTATTCTTATGTGACAGTTGAAATTTTGGATCAATTAGAGAAAAATACTGGCGCCA   1766

Qy   1141 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1767 TTAACGATACAAGATTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT   1826

Qy   1201 ATTAATGTTATTAGAGACAATTGATAATGCGTTGATTTTATTTATGGGTGTGCAGTGATATT   1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 ATTAATGTTATTAGAGACAATTGATAATGCGTTGATTTTATTTATGGGTGTGCAGTGATATT   1886

Qy   1261 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTTAA   1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1887 CAAAAATTACAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTTAA   1940
```

RESULT 2
US-09-134-001C-2296
; Sequence 2296, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCU
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2296
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2296

Query Match   70.0%;   Score 922; DB 4; Length 1329;
Best Local Similarity 81.4%; Pred. No. 5.1e-205;
Matches 1069; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

```
Qy      1 ATGACACCATCATCATTTGTCGGCATGAGTTCTGGCATGAGTTCATTAGCACACAAATC     60
          ||||||||| || ||||||||||   |||  || ||||||||||||||||||||||||
Db     16 ATGACACACTATCATCATTTGTCGGAATTAAAGGCTCAGGCATGAGTTCATTAGCACACAAATC     75

Qy     61 ATGCATGATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTTACAGAA    120
          ||||||| |||||||||||||||||||||||||||||||||||||| |||||||||| |||||
Db     76 ATGCATGACGTCGGTCATGAAGTCAAGTAGTACAAGGTTCAGACATAGAATCATCACGTTTTTTACAGAA    135

Qy    121 GTTGCTCTTAGAAATAAGGGGATAAAAAATATTACCATTGATGCTAATAACATAAAGAA    180
          |||||||  || ||||||| ||| ||||| ||||||| |||| |||||||| || ||||
Db    136 GTTGCCATTAAGAAGAAATAAAAGGGATTAAAATTTTTACCTTTTGATGCAATAATATTATTACAAAA    195

Qy    181 GATATGGTAGTTATACAAGTAATGTCATTCGCGAGTAGCCATGAGAGAAATAGTACGTGCA    240
          ||||||||||| |||||||||||  ||||| |||| ||||||||| || ||||| ||||||
Db    196 GATATGGTAGTTATTACAAGTAATGTAATCATGAAATCATGAAGAAAATTGTTAGGGGCA    255

Qy    241 CATCAATTGAAAATTGATGTTGTAAGTTAGTGAAGTTATTATATAAATATTTTAGGACAGATTATTGATCAA    300
          ||||||||||  ||||||||| || ||| || || ||||||||| ||||| ||||||||| ||
Db    256 CATGAATTGAAATTGATGTGTTGATTATTTTTAGGACGATATTGTTCTTTGGTTCATGTCTATTTTTAGGACAGATCAA    315

Qy    301 TATACTTCAGTAGCTGTAACTGGTGTACTGGTGCACATGGTAAAACTTCTACACAGGTTATTATCA    360
          |||| |  ||  || | ||| ||| | |||| | ||| |||| |||||| |||| || || |||||
Db    316 TATACATCCGTTGCTGTCGTGTTACTGGTAACGGTGTAACAGTTTATTATCA    375

Qy    361 CATGTTATGAATGGTGATAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA    420
          |||||||||||||||||||||||||||||||||| | |||| |||||| |||||||||
Db    376 CATGTTGATGAATGGTGATAAAAGAACATCATTCCTAATTCCTAATCCTGATAAAAAAGACTTCATTTTTAATTGGTGATGGTACGGTATGGGA    435
```

APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 619 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-179

Query Match   43.3%;   Score 570.4;   DB 4;   Length 619;
Best Local Similarity 98.0%;   Pred. No. 1.6e-123;
Matches 588; Conservative   0; Mismatches 11; Indels   1; Gaps   1;

Qy   716 AAGGTACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCAC 775
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   616 ACGAGACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCAC 557

Qy   776 AATATGTGACCATACAGTTTAAATGCATTGTTAAATGCATTAGTATTAGTATTATTAGAGA 835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   556 AATATGTGACCATACAGTTTAAATGCATTGTTAAATGCATTAGTATTAGTATTATTAGAGA 497

Qy   836 AGCTAGATGTTACAAATTACAAATTAAAGAAGCATTAGAAACGTTTGGGTGTGTTAAACGTTCGTT 895
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496 AGCTAGATGTTACAAATTACAAATTAAAGAAGCATTAGAAACGTTTGGGTGTGTTAAACGTTCGTT 437

Qy   896 TCAATGAAACTACAATTGCAAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAG 955
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   436 TCAATGAAACTACAATTGCAAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAG 377

Qy   956 AAATTAGTGCTACAATTGACAATTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAG 1015
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376 AAATTAGTGCTACAATTGACAATTGACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAG 317

Qy 1016 TATTTCAACCACACTTCTCTAGAACACAGGCATTTTTAAATGAATTCGAGAAAGTT 1075
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   316 TATTTCAACCACACTTCTCTAGAACACAGGCATTTTTAAATGAATTCGAGAAAGTT 257

Qy 1076 TAAGTAAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTG 1135
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    256  TATGTAAGCAGATCGTCGTCTATTCTTATGTGAAAATTTTTGGCTCAATTAGAGAAAATTCTG  197

Db    190  GTTATACAAGGTAAGCGATT---CGGCGAGTAGCCATGAAGAGAAATAGTACGTGCACATCAA  246
           ||||||
Qy   1569  ATTATCGCTGGAAATGCCTTTCGTCCAGATAACAACGTCGAAATTGCCTATGCGGACCAA    1510

Qy   1136  GCGCATTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAG   1195
Qy    247  TTGAAATTAGATGTGTGTAAGTATAATGATTTTTAGGACAGATTATTGATCAATATACT    306
Db    196  GCGCCATTAACGATACAAGATTTAATTGATAAAATTGGGAGTGCATCGTTCATTAATGAAG   137
Db   1509  AATGGTATCAGCTACAAACGTTACCATGAGTTTCTAGTAGCTTTATGCGTGACTTTGTT   1450

Qy   1196  ATTC-TATTAATGTATTAGAACAATTTGATAATGCTGTGTTTATTTTATTTATGGGTGCAGGT  1254
Db    136  ATTCTTATTAATGTATTAGACAATTTGATAATGCTGTGTTGTTTTATTTATGGGTGCAGGT    77
Qy    307  TCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACACAGTTTATTATCACATGTT    366

Qy   1255  GATATTCAAAAATTACAAAATGCATATTTGAGATAAAATTAGGCATGAAAAATGCGTTTAA  1314
Db   1449  AGCATGGGAGTAGCAGGAGCACATGAAAAACTTCAACGACAGTATGTTGTCTCATGTC   1390
Db     76  GATATTCAAAAATTACAAAAATGCATATTTAGATAAATTAGGCATGAAAAATGCGTTTTAA    17
Qy    367  ATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGATTGCCT   426

Db   1389  TTGTCTCCACATTACAGATACCAGCTTTCTTGATTGGAGATGGAGCAGGTCGTGGTTCGGCC  1330
Qy    427  GAAAGTGATTATTCGCTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGTTATAAA    486
Db   1329  AATGCCAAATATTTTGTCTTTGAATCTGACGAATATGAGCGTCACTCCATGCCTTACCAC   1270

Qy    487  CCTGATTACGCAATTATGACACAATATATGATTTCGATCATCCTGATTATTTTAAAGATATT  546
Db   1269  CCAGAATACTCTATTATCACCAACATTGACTTTGACCATCCAGATTATTTCACAAGTCTC   1210
Qy    547  AATGATGTTTTGATGCATTCCAAGAAATGGCACATATATGTTAAAAAAGGTATTATTGCT   606

Db   1209  GAGGATGTTTTTTAAATGCCTTTAACGACTATGCCAAACAAATCACCAGGGTCTTTTGTC   1150
Qy    607  TGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTTATTATTATGA    666
Db   1149  TATGGTGAAGATGCTGAATTGCGAATTGCGTAACATTACGTCTGATGCACCAATTTATTATTGT   1090

Qy    667  TTTAAAGATTCGG---ATGACATTTATGCTCAAAATATTCAAAATTACGGATAAAGTACT    723
Db   1089  TTTGAAGCTGAAGGCAATGACTTTGTAGCTAGTGATCTTCTTCGTTCAATAACTGGTTCA   1030
Qy    724  GCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCGTGTCTCCACAATATGGT    783

Db   1029  ACCTTCACCGTTCATTTCCGTGGACAAAACTTGGGCAATTCCACATTCCACCTTTTGGT   970
Qy    784  GACCATACAAGTTTAAATGCATTAGCTGTGAATTGCGATTAGTTATTTAGAGAAGCTAGAT    843
Db    969  CGTCACAATATCATGAAATGCGACAGCGTATTGGTCTTCTTTACACAGCAGGATTTGAT    910

Qy    844  GTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGTTGGTCGTTAAACGTCGTTTCAATGAA   903
Db    909  TTGAACTTGGTGTGCGTGAGCACTTGAAAAACATTTGCCGGTGTTAAACGTCGTTTCACTGAG   850
Qy    904  ACTACAAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAGAAATTAGT   963

Db    849  AAAAATTGTCAATGATACAGTGATTATCGATGACTTTGCCCACCATCCAACAGAAATTATT   790
Qy    964  GCTACAATTGAAACAGCACGACGAAAGAAAATATCCACATAAAGAAGTGTTGCAGTATTTCAA  1023
Db    789  GCGACCTTGGATGCGGCTCGTCGTCAGAAAATCACCCAAGCAAGAAATTGTAGCAGTCTTTCAA   730

Qy   1024  CCACACACTTCTCTAGAACACACGGCATTTTAAATGAATTGCAGAAAGTTAAGTAAGTTAAGTAA  1083
Db    729  CCGCATACCTTTACAAGACCATTGCCTTGTTGGACGACTTGCCCATGCTTTAAACCAA    670
Qy   1084  GCAGATCGTGTATTCTTATGTGAAAATTTTGGATCAATTAG--AGAAATACTGCGCA  1140

Db    669  GCAGATGCGTTTATCTACCGCAAATTTATGGCCTTCGTGAAGTAGATCATGGTGAC    610
Qy   1141  TTAACGATACAAGATTTAATTGATAAAAT---TGAAGGTGCATCGTTAATTAATGAGAT  1197
Db    609  GTTAAGGTAGAAGACCTTAGCCAACAAATCAACAAAAAACAACCAAGTAGTACTTGTTGAA   550

Qy   1198  TCTATTAATGTATTAGACAAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGAT   1257
Db    549  AATGTTTCTCCACTCCTAGAACCATGACCATGACAATGCTGTTTACGTCTTTTACGTCGATTATGATGAA    490
Qy   1258  ATTCAAAAATTACAAAATGCATAT  1281
```

```
  Query Match        30.4%;      Score 400.8;    DB 4;   Length 11864;
  Best Local Similarity  59.2%;  Pred. No. 9.3e-84;
  Matches   760;  Conservative    0;  Mismatches  512;  Indels  12;  Gaps  4;

Qy     10  TATCATTTGTCGGAATTAAAGGTTCTGGCATGAAGTTCATTAGCACAAATCATGCATGAT    69
Db   1749  TATCATTTTATCGGAATTAAGGGATCAGGGATCATCAGGATGCCTTGGCCTTGATGTTGCACCAG   1690
Qy     70  TTAGGACATGAAGTTCAAGGATTCAAGGATCCGGATATTCGATCACGTATTACGAGAGTTGCTCTT   129
Db   1689  ATGGGGCACAAGGTTCAGGGATCATCAGGATGTTGAAAAGTTGAAAAGTTACTACTTTACCCACGCGGTCTT   1630
Qy    130  AGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAGAAGATATGGTA   189
Db   1629  GAGCAGGCAGGAGGAATTACCACATTCTTCCCTTTTTGATGAAAAAAATCTAGACCGGTGATATGGAA   1570
```

```
Db  489 ...ATCCAAACCTATGAATACTTCATTT 466

Db  588 ATTATCGCTGGAAATGCCTTCGTCCAGATAACAACGTCGAAATTGCCTATGCGGACCAA 647
Qy  247 TTGAAATTAGATGTGTAAGTATAATGATTAGGACAGATTATTGATCAATATACT 306
Db  648 AATGGTATCAGTCAGACTGTACCATGAGTTCTAGGTAGCTTATGCGTGACTTTGTT 707
Qy  307 TCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCACATGT 366
Db  708 AGCATGGGAGTAGCAGGAGGCACATGCAGAAAACTTCAACGACAGGGTATGTGTCTCATGTC 767
Qy  367 ATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGATTGCCT 426
Db  768 TTGTCTCACATTACAGATACCAGCTTCTTGATTGGAGATGGGACAGGTCGTGGTTCGGCC 827
Qy  427 GAAAGTGATTATTCGCTTTGAGGCATGTGACGAATATAGACGGTCACTTTTAGTTATAAA 486
Db  828 AATGCCAAATATTTTGTCTTTGAATCTGACATTGAGCGTCACTCATGCCTTACCAC 887
Qy  487 CCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTAAAGATATT 546
Db  888 CCAGAATACTCTATTATCACCAACATTGACTTGACCATCCAGATTATTTCACAAGTCTC 947
Qy  547 AATGATGTTTTTGATGCATTCCAAGAAATGCGACATAATGTTAAAAAGGTATTATTGCT 606
Db  948 GAGGATGTTTTCAATGCCTTTAACGACTATGCCAAACAAATTACCAAGGGTCTTTTTGTC 1007
Qy  607 TGGGGTGATGATGAACATCGATAAATTGAAGCAGATGTTCCAATTATTATTATTATTGGA 666
Db 1008 TATGGTGAAGATGCTGAAATTGCGTAAGATTACGTCTGATGCACCAATTTATTATTATGGT 1067
Qy  667 TTTAAAGATTCGG---ATGACATTTATGCTCAAAATAATTCAAATTACGGATAAAGGTACT 723
Db 1068 TTTGAAGCTGAAGGCAATGACTTTGTGAGCTAGTAGTTTTATGATCACTTCCTGTCTCCACAATATGGT 1127
Qy  724 GCTTTTGATGTGTATGTGGATGGTGAGGTTTTATGATCACTTCCTGTCTCCACAATATGGT 783
Db 1128 ACCTTCACCGTTCATTTCCGTGGACAAAACTTGGGGCAATTCCACATTCCACCCTTTGGT 1187
Qy  784 GACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTAGTTATTTAGAAGCTAGAT 843
Db 1188 CGTCACAATATCATGAATGCGACAGCCGGTTATTGGTCTTCTTTACACAGCAGGATTTGAT 1247
Qy  844 GTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAATGAA 903
Db 1248 TTGAACTTGGTGCGTGAGCACTTGAAAAACATTGCCGGTGTTAAACGTCGTTTCACTGAG 1307
Qy  904 ACTACAATTGCAAATCAAGTTATTGTGATGATTATGTGTAGATGATTATGCACCATCCAAGAAATTAGT 963
Db 1308 AAAATTGTCAATGATACAGTGATTTATTGATGACTTGCCCATCATCCAACAAGAAATTATT 1367
Qy  964 GCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTTCAA 1023
Db 1368 GCGACCTTGGATGCGGCGCTCGTCAGAAAATACCCAGCAAGCAAGGAAAATTGTAGCAGTCTTTCAA 1427
Qy 1024 CCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGTAAA 1083
Db 1428 CCGCATACACTTTTACAAGAACCATTGCCCGTTGGACGACTTTGCCCATGCTTTAAACCAA 1487
Qy 1084 GCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAG---AGAAAATACTGGCGCA 1140
Db 1488 GCAGATGCTGTTTATCTAGCGCAAATTTATGCGCTCGGCTCGTGAAGTAGATCATGGTGAC 1547
Qy 1141 TTAACGATACAAGATTTAATTGATAAAAT---TGAAGGTGCATCGTTAATTAATGAAGAT 1197
Db 1548 GTTAAGGTAGAGACCTAGCCAATAAAATCAACAAAAACACCAAGTGATTACTGTTGAA 1607
Qy 1198 TCTATTAATGTATTAGGAGACAATTTGATAATGACTGTGTTATTTATTATTATGGTGCAGGTGAT 1257
Db 1608 AATGTTTCTCCACTCCTAGACCATGACCAATGCGTGTTTTACGTCTTTATGGGAGCAGGAGAC 1667
Qy 1258 ATTCAAAAATTACAAAATGCATAT 1281
Db 1668 ATCCAAACCTATGAATACTTCATTT 1691
```

```
Query Match    30.3%;  Score 399.2; DB 4;  Length 2715;
Best Local Similarity   59.1%;  Pred. No. 1.4e-83;
Matches 759;  Conservative   0;  Mismatches 513;  Indels  12;  Gaps  4;

Qy   10 TATCCATTTGTGCGGAATTAAAGGTTCTCGGCATAGTAAGTTCATTAGCACAAATCATGCATGCATGAT 69
Db  408 TATCCATTTATCGGAATTAAGGGATCAGGGATGAGTGGCCTTGGCCTTTGATGTGTTGCACCAG 467
Qy   70 TTAGGACATGGAAGTTCAAGGATCGGATATTGAGAACTAGTATTTTACAGAAGTTGCTCTT 129
Db  468 ATGGGGCACAAGGTTCAGGGATCAGGATGTGTTGAAAAGTTGAAAAGTACTTTACCCAACGCGGGTCTT 527
Qy  130 AGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAAGATATGGTA 189
Db  528 GAGCAGGCAGGAATTACCATTCTTCCTTTGATGAAAAGAAATCAGACGGTGATATGGAA 587
Qy  190 GTTATACAAGGTAATGCATT--CGCGAGTAGCCATGCAGAGAATGAAGAAATAGTACGTGCACATCAA 246
```

Db 711 TGTTGCAGTAINTCAAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTG 652

Qy 1067 CAGAAAGTTTAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTGGATCAATTAG 1124
Db 651 CAGAAAGTTTAAGTAAAGCAGATCGTGTATTTCTTATTGTGAAATTTTGGATCAATTAG 592

Qy 1125 AGAAATACTGGCGCCATTAACGATACAAGAGATTTAATTGATCAAGATTGAAGGTGCATCGTT 1184
Db 591 AGAAATACTGGCGCCATTAACGATACAAGAGATTTAATTGATAAAAATTGAAGGTGCATCGTT 532

Qy 1185 AATTAATGAAGATTCTATTAATGTATTAGAACAATTGATAATGCTGTTATTTT-ATTTA 1243
Db 531 AATTAATGAAGATTCTATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTCATTTA 472

Qy 1244 TGGGTGCAGGTGATATTCAAAAATTACAAATGCATATTTAGATAAAATAGGCATGAAAA 1303
Db 471 TGGGTGCAGGTGATATTCAAAAATGCATATTTAGATAAAATAGGCATGAAAA 412

Qy 1304 ATGCGTTTTAA 1314
Db 411 ATGCGTTTTAA 401

RESULT 7
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,315
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:

RESULT 6
US-08-714-918-61/c
; Sequence 61, Application US/08714918
; Patent No. 6037123
GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,918
; FILING DATE: September 13, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-714-918-61

Query Match   29.6%; Score 389.6; DB 3; Length 888;
Best Local Similarity 94.1%; Pred. No. 1.7e-81;
Matches   462; Conservative   1; Mismatches 21; Indels   7; Gaps   6;

Qy 828 TTTAGAGAAGCTAGATGTTACAAATATTAAAGAAGCATTAAAGAAACGTTGGTGGTGTTAA 887
Db 888 TTTAGAGAAGTTAGATG-TCCCAATNTTAAAGAAGCATTAAAGAAACG-TTGGTGGNGTTAA 831

Qy 888 ACGTCGTTCAATGAAACTACAATTGCAAATCAGTTATTGTGATGATTATGCACACCA 947
Db 830 ACGTC-NNTCNATGANACTACAATCGCAAATCAGTTAATGTGAGCTGATTATGNACACCA 772

Qy 948 TCCAAGAGAAATTAGTGCTACAAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGT 1007
Db 771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCGCACGAAAGAAATATCCACATAAAGAAGT 712

Qy 1008 TGTTGCAGTATTTC-AACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTG 1066

;   LENGTH: 888 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-265-315-61

Query Match   29.6%;   Score 389.6; DB 3; Length 888;
Best Local Similarity   94.1%;   Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

```
Qy   828 TTTAGAGAAGCTAGATGTTACAAATATTAAAGAAGAGAGCATTAGAAACGTTTGGTGGTGTTAA 887
         |||||||||| ||||| ||  |||||||||||||||||| ||||||||||| ||| |||||||
Db   888 TTTAGAGAAGTTAGATG-TCCCAATNTTAAAGAAGAGCATTAGAAACG-TTGGTGGNGTTAA 831

Qy   888 ACGTCGTTTCAATGAAACTACAAGTTGCAAATCAAGTTGTAGATGATTATGCACACCA 947
         ||||| |  |  |||| |||||||  |||||||||||||  |||||  ||| |||||
Db   830 ACGTC-NNTCNATGANACTACAATCGCAAATCAAGTTAATGTAGCTGATTATGNACACCA 772

Qy   948 TCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAGAAGAAATATCCACATAAAGAAGT 1007
         ||||||||| ||||||||  |||||||||| ||||||||||||||||||||||||||||
Db   771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCGCACGAGAAGAAATATCCACATAAAGAAGT 712

Qy  1008 TGTTGCAGTATTTC-AACCACACACTTTCTCTAGAACACACACTTTTAAATGAATTTG 1066
         ||||||||||| || ||||||| ||||||||||||||||||||||||||||||||||
Db   711 TGTTGCAGTATNTCAAACCACACCACACACTTTCTCTAGAACACACACTTTTAAATGAATTTG 652

Qy  1067 CAGAAAGTTTAAGTAAAGCAGCAGATCGTGTGATT--CTTATGTGAAATTTTTGGATCAATTAG 1124
         |||||||||||||||||||||||||||||||||    |  ||||||| |||||||||||||||
Db   651 CAGAAAGTTTAAGTAAAGCAGCAGATCGTGTATTTCTTATTCTGAAATTTGTGAAATTTTTGGATCAATTAG 592

Qy  1125 AGAAAATACTCGGCGCATTAACGATACAAGATTTAATTGAAATTGATAAAAATTGAAGGTGCATCGTT 1184
         |||||||||||||||||||||||| |||||||||||||   |||:|||||||||||||||||||||
Db   591 AGAAAATACTCGGCGCATTAACGAWACAAGATTTAATTAATGTATTGATAAAAATTGAAGGTGCATCGTT 532

Qy  1185 AATTAATGAAGATTCTATTAATTGTATTAGAACAATTTGATAATGCTGTTATTTT-ATTTA 1243
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||||
Db   531 AATTAATGAAGATTCTATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTTCATTTA 472

Qy  1244 TGGGTGCAGGTGATATTCAAAAATTACAAAATTAGATAAAATTAGGCATGAAAA 1303
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471 TGGGTGCAGGTGATATTCAAAAATTACAAAATTAGATAAAATTAGGCATGAAAA 412

Qy  1304 ATGCGTTTTAA 1314
         |||||||||||
Db   411 ATGCGTTTTAA 401
```

RESULT 8
US-09-265-315-61/c
; Sequence 61, Application US/09265315
; Patent No. 6187541
GENERAL INFORMATION:
;   APPLICANT: Benton, Bret
;   APPLICANT: Lee, Ving J.
;   APPLICANT: Malouin, Francois
;   APPLICANT: Martin, Patrick K.
;   APPLICANT: Schmid, Molly B.
;   APPLICANT: Sun, Dongxu
;   TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
;   TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
;   TITLE OF INVENTION: TARGET GENES
;   NUMBER OF SEQUENCES: 111
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Lyon & Lyon
;   STREET: 633 West Fifth Street
;   STREET: Suite 4700
;   CITY: Los Angeles
;   STATE: California
;   COUNTRY: U.S.A.
;   ZIP: 90071-2066
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;   MEDIUM TYPE: storage

;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: IBM P.C. DOS 5.0
;   SOFTWARE: Word Perfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/265,315
;   FILING DATE: March 9, 1999
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/714,918
;   FILING DATE: September 13, 1996
;   APPLICATION NUMBER: 60/009,102
;   FILING DATE: December 22, 1995
;   APPLICATION NUMBER: 60/003,798
;   FILING DATE: September 15, 1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Warburg, Richard J.
;   REGISTRATION NUMBER: 32,327
;   REFERENCE/DOCKET NUMBER: 240/247
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (213) 489-1600
;   TELEFAX: (213) 955-0440
;   TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 888 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-265-315-61

Query Match   29.6%;   Score 389.6; DB 3; Length 888;
Best Local Similarity   94.1%;   Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

RESULT 9
US-09-266-417-61/c

Qy 1067 CAGAAAGTTTAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTTGGATCAATTAG 1124
         ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db 651 CAGAAAGTTTAAGTAAAGCAGATCGTGTATTCTTATTGTGAAATTTTTGGATCAATTAG 592

Qy 1125 AGAAAATACTGGCGCATTAACGATACAAGAGATTTAATTGATAAAATTGAGGTGCATCGTT 1184
         ||||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db 591 AGAAAATACTGGCGCCATTAACGAWACAAGAGATTTAATTGATAAAATTGAGGTGCATCGTT 591

Qy 1185 AATTAATGAAGATTCATTAATTGTATTAGAACAATTGATAATGCTGTTATTT--ATTTA 1243
         |||||||||||||| |||| ||| |||||||||||||||||||||||||||| |||||
Db 531 AATTAATGAAGATTCTATTATTAGAACAATTGATAATGCTGTTATTTTCATTTA 472

Qy 1244 TGGGTGCAGTGATATTCAAAATTCAAAATTACAAAATGCATATTTAGATAAAATTAGGCATGAAAA 1303
         |||||||||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db 471 TGGGTGCAGTGATATTCAAAAATTCAAAAATTCATATTTAGATAAAATTAGGCATGAAAA 471

Qy 1304 ATGCGTTTTAA 1314
         |||||||||||
Db 411 ATGCGTTTTAA 401

RESULT 10
US-09-528-709-61/c
; Sequence 61, Application US/09528709
; Patent No. 6630303
GENERAL INFORMATION:
   APPLICANT: Benton, Bret
              Lee, Ving
              Malouin, Francois
              Martin, Patrick K.
              Schmid, Molly B.
              Sun, Dongxu
   TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
                       TARGET GENES
   NUMBER OF SEQUENCES: 111
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Lyon & Lyon
      STREET: 633 West Fifth Street
              Suite 4700
      CITY: Los Angeles
      STATE: California
      COUNTRY: U.S.A.
      ZIP: 90071-2066
   COMPUTER READABLE FORM:
      MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
                   storage
      COMPUTER: IBM Compatible
      OPERATING SYSTEM: IBM P.C. DOS 5.0
      SOFTWARE: Word Perfect 5.1
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/528,709
      FILING DATE: 17-Mar-2000
      CLASSIFICATION: 424
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/714,918
      FILING DATE: September 13, 1996
      APPLICATION NUMBER: 60/009,102
      FILING DATE: December 22, 1995
      APPLICATION NUMBER: 60/003,798
      FILING DATE: September 15, 1995
   ATTORNEY/AGENT INFORMATION:
      NAME: Warburg, Richard J.
      REGISTRATION NUMBER: 32,327
      REFERENCE/DOCKET NUMBER: 222/005
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (213) 489-1600
      TELEFAX: (213) 955-0440
      TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 888 base pairs
      TYPE: nucleic acid

; Sequence 61, Application US/09266417
; Patent No. 6228588
GENERAL INFORMATION:
   APPLICANT: Benton, Bret
              Lee, Ving J.
              Malouin, Francois
              Martin, Patrick K.
              Schmid, Molly B.
              Sun, Dongxu
   TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
                       ACTIVE ON STAPHYLOCOCCUS AUREUS
   TITLE OF INVENTION: TARGET GENES
   NUMBER OF SEQUENCES: 111
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Lyon & Lyon
      STREET: 633 West Fifth Street
      STREET: Suite 4700
      CITY: Los Angeles
      STATE: California
      COUNTRY: U.S.A.
      ZIP: 90071-2066
   COMPUTER READABLE FORM:
      MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
      MEDIUM TYPE: storage
      COMPUTER: IBM Compatible
      OPERATING SYSTEM: IBM P.C. DOS 5.0
      SOFTWARE: Word Perfect 5.1
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/266,417
      FILING DATE: March 9, 1999
      CLASSIFICATION: 435
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 08/714,918
      FILING DATE: September 13, 1996
      APPLICATION NUMBER: 60/009,102
      FILING DATE: December 22, 1995
      APPLICATION NUMBER: 60/003,798
      FILING DATE: September 15, 1995
   ATTORNEY/AGENT INFORMATION:
      NAME: Warburg, Richard J.
      REGISTRATION NUMBER: 32,327
      REFERENCE/DOCKET NUMBER: 240/248
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (213) 489-1600
      TELEFAX: (213) 955-0440
      TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 888 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
US-09-266-417-61

Query Match   29.6%;   Score 389.6; DB 3;   Length 888;
Best Local Similarity 94.1%;   Pred. No. 1.7e-81;
Matches 462; Conservative 1; Mismatches 21; Indels 7; Gaps 6;

Qy 828 TTTAGAGAAGTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAA 887
         ||||||||||||||| |||||| |||||||||||||||||||| ||||||| ||||||
Db 888 TTTAGAGAAGTAGATG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGNGNGTTAA 831

Qy 888 ACGTCGTTTCAATGAAACTACAAGTCGAAATCAAGTTATTGTGATGATTATGCACACCA 947
         ||||| || ||| | ||||||| ||||||||||||| ||| | |||||||| ||||||
Db 830 ACGTC-NNTCATGANACTACAATCGGCAAATCAAGTTAATGTAGCTGATTATGNACACCA 772

Qy 948 TCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGT 1007
         |||||||||| ||||||| |||||||||| ||||||||||||||||||||||||||||
Db 771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCCACGAAAGAAATATCCACATAAAGAAGT 712

Qy 1008 TGTTGCAGTATTC-AACCACACACTTTCTCAGAACACAGGCATTTTAAATGAATTTG 1066
         ||||||||||| | |||||||||||||||||||||||||||||||||||||||||||
Db 711 TGTTGCAGTATNTCAAACCACACACTTTCTCAGAACACACAGGCATTTTAAATGAATTTG 652

;
;
;   STRANDEDNESS: single
         TOPOLOGY: linear
         SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-528-709-61

Query Match   29.6%;   Score 389.6; DB 4;   Length 888;
Best Local Similarity   94.1%;   Pred. No. 1.7e-81;
Matches 462; Conservative   1; Mismatches 21; Indels   7; Gaps   6;

Qy   828 TTTAGAGAAGCTAGATGTTACAAATATTAAAGAGAGCATTAGAAACGTTTGGTGGTGTTAA 887
          ||||||||||||||||||||| ||||| |||||||||||||||||||||||| || |||||
Db   888 TTTAGAGAAGTTAGATG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGGGNGTTAA 831

Qy   888 ACGTCGTTTCAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCA 947
          ||||| ||| ||||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   830 ACGTC-NNTCNATGANACTACAATCGCAAATCAAGTTAATGTAGCTGATTATGNACACCA 772

Qy   948 TCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGT 1007
          |||||||||| |||||||||||||| ||||||| |||||||||||||||||||||||||
Db   771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCGACGAAAGAAATATCCACATAAAGAAGT 712

Qy   1008 TGTTGCAGTATTTC-AACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 1066
          |||||||||||||| ||||||| |||||||||||||| |||||||||||||||||||||
Db   711 TGTTGCAGTAINTCAAACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 652

Qy   1067 CAGAAAGTTTAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTTGGATCAATTAG 1124
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   651 CAGAAAGTTTAAGTAAAGCAGATCGTGTATTTCTTATTGTGAAATTTTTGGATCAATTAG 592

Qy   1125 AGAAAATACTGGGCCATTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTT 1184
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   591 AGAAAATACTGGGCGCATTAACGAWACAAGATTTAATTGATAAAATTGAAGGTGCATCGTT 532

Qy   1185 AATTAATGAAGATTCATTATTGTATTAGAACAATTTGATAAATGCTGTTATTTT-ATTTA 1243
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db   531 AATTAATGAAGATTCATTATTGTATTAGAACAATTTGATAATGCTGTTATTTTTCATTTA 472

Qy   1244 TGGGTGCAGGTGATATTCAAAAATTACAAAAATTACAAAAATGCATATTTAGATAAATGAAAA 1303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471 TGGGTGCAGGTGATATTCAAAAATTACAAAAATGCATATTTAGATAAAATTAGGCATGAAAA 412

Qy   1304 ATGCGTTTTAA 1314
          |||||||||||
Db   411 ATGCGTTTTAA 401

RESULT 11
US-09-527-745-61/c
; Sequence 61, Application US/09527745
; Patent No. 6638718
GENERAL INFORMATION:
     APPLICANT: Benton, Bret
               Lee, Ving
               Malouin, Francois
               Martin, Patrick K.
               Schmid, Molly B.
               Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
                   TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
     ADDRESSEE: Lyon & Lyon
     STREET: 633 West Fifth Street
             Suite 4700
     CITY: Los Angeles
     STATE: California
     COUNTRY: U.S.A.
     ZIP: 90071-2066
COMPUTER READABLE FORM:
     MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
                  storage
     COMPUTER: IBM Compatible
     OPERATING SYSTEM: IBM P.C. DOS 5.0

;
;
;
          SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/09/527,745
     FILING DATE: 17-Mar-2000
     CLASSIFICATION: 424
PRIOR APPLICATION DATA:
     APPLICATION NUMBER: US/08/714,918
     FILING DATE: September 13, 1996
     APPLICATION NUMBER: 60/009,102
     FILING DATE: December 22, 1995
     APPLICATION NUMBER: 60/003,798
     FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
     NAME: Warburg, Richard J.
     REGISTRATION NUMBER: 32,327
     REFERENCE/DOCKET NUMBER: 222/005
TELECOMMUNICATION INFORMATION:
     TELEPHONE: (213) 489-1600
     TELEFAX: (213) 955-0440
     TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
     LENGTH: 888 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
     SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-527-745-61

Query Match   29.6%;   Score 389.6; DB 4;   Length 888;
Best Local Similarity   94.1%;   Pred. No. 1.7e-81;
Matches 462; Conservative   1; Mismatches 21; Indels   7; Gaps   6;

Qy   828 TTTAGAGAAGCTAGATGTTACAAATATTAAAGAGAGCATTAGAAACGTTTGGTGGTGTTAA 887
          ||||||||||||||||||||| ||||| |||||||||||||||||||||||| || |||||
Db   888 TTTAGAGAAGTTAGATG-TCCCAATNTTAAAGAAGCATTAGAAACG-TTGGTGGGNGTTAA 831

Qy   888 ACGTCGTTTCAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCA 947
          ||||| ||| ||||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   830 ACGTC-NNTCNATGANACTACAATCGCAAATCAAGTTAATGTAGCTGATTATGNACACCA 772

Qy   948 TCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGT 1007
          |||||||||| |||||||||||||| ||||||| |||||||||||||||||||||||||
Db   771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCGACGAAAGAAATATCCACATAAAGAAGT 712

Qy   1008 TGTTGCAGTATTTC-AACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 1066
          |||||||||||||| ||||||| |||||||||||||| |||||||||||||||||||||
Db   711 TGTTGCAGTAINTCAAACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 652

Qy   1067 CAGAAAGTTTAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTTGGATCAATTAG 1124
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   651 CAGAAAGTTTAAGTAAAGCAGATCGTGTATTTCTTATTGTGAAATTTTTGGATCAATTAG 592

Qy   888 ACGTCGTTTCAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCA 947
          ||||| ||| ||||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   830 ACGTC-NNTCNATGANACTACAATCGCAAATCAAGTTAATGTAGCTGATTATGNACACCA 772

Qy   948 TCCAAGAGAAATTAGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGT 1007
          |||||||||| |||||||||||||| ||||||| |||||||||||||||||||||||||
Db   771 TCCAAGAGAANTTAGTGCNNCAATTGAAACCGACGAAAGAAATATCCACATAAAGAAGT 712

Qy   1008 TGTTGCAGTATTTC-AACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 1066
          |||||||||||||| ||||||| |||||||||||||| |||||||||||||||||||||
Db   711 TGTTGCAGTAINTCAAACCACACACTTTCTCTAGAACACACAGCCATTTTTAAATGAATTTG 652

Qy   1067 CAGAAAGTTTAAGTAAAGCAGATCGTGTATT--CTTATGTGAAATTTTTGGATCAATTAG 1124
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   651 CAGAAAGTTTAAGTAAAGCAGATCGTGTATTTCTTATTGTGAAATTTTTGGATCAATTAG 592

Qy   1125 AGAAAATACTGGGCCATTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTT 1184
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   591 AGAAAATACTGGGCGCATTAACGAWACAAGATTTAATTGATAAAATTGAAGGTGCATCGTT 532

Qy   1185 AATTAATGAAGATTCATTATTGTATTAGAACAATTTGATAAATGCTGTTATTTT-ATTTA 1243
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db   531 AATTAATGAAGATTCATTATTGTATTAGAACAATTTGATAATGCTGTTATTTTTCATTTA 472

Qy   1244 TGGGTGCAGGTGATATTCAAAAATTACAAAAATTAGGCATGAAAA 1303
          ||||||||||||||||||||||||||||||||||||||||||||
Db   471 TGGGTGCAGGTGATATTCAAAAATTACAAAAATGCATATTTAGATAAAATTAGGCATGAAAA 412

Qy   1304 ATGCGTTTTAA 1314
          |||||||||||
Db   411 ATGCGTTTTAA 401

RESULT 12
US-08-961-083-115
; Sequence 115, Application US/08961083

```
; Patent No. 6159469
; GENERAL INFORMATION:
;  APPLICANT: Choi et. al.
;  TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;  NUMBER OF SEQUENCES: 452
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,083
;   FILING DATE:
;   CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P2
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 115:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 1267 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
US-08-961-083-115
```

```
Query Match       28.1%;   Score 370.2;   DB 3;   Length 1267;
Best Local Similarity 58.6%;   Pred. No. 5.9e-77;
Matches 721;   Conservative  0;   Mismatches 498;   Indels 12;   Gaps 4;
```

```
Qy   63 GCATGATTAGGACATCGAAGTTCAAGGATCGGATCGGATATTGAGAACTACGTATTTACAGAAGT 122
           |||   ||||||  |||  || |||||||  ||||||| |||
Db    1 GCACCAGATGGGCACAAGGTTCAGGGATCAGGATCGATGTTGAAAGTACTACTTTACCCAACG   60

Qy  123 TGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATACATAAAGAGAGA 182
           |||||          |||  ||  |||||||
Db   61 CGGTCTTGAGCAGCAGGAGGAATTACCATTCTCCTTTGATGAAAAAAATCTAGACGGTGA 120

Qy  183 TATGGTAGTTATACAAGGTAATGCATT---CGCGAGTAGCCATGAAGAAATAGTTACGTGC 239
           |||||            |||  ||||   ||  ||
Db  121 TATGGAAATTATCGCTGGAAATGCCTTTCGTCCAGATAAACACGTCGAAATTGCCATGC 180

Qy  240 ACATCAATTGAAATTGAAATTAGATGTTGTAAGTTATTAAATGATTTTTAGGACAGATTATTGATCA 299
           ||                |||||||||
Db  181 GGACCAAAATGGTATCAGCTACAAACGTTACCATGAGTTTCTGGTAGCTTTATGCGTGA 240

Qy  300 ATATACTTCAGTAGCTGTGTAACTGGTGCACATGGTAAACTTCTACAACAGGTTTATTATC 359
           ||    |||  ||| ||  ||||  ||||||| || ||||||
Db  241 CTTTGTTAGCATGGAGTAGCAGGAGCACATGGAAAAACTTCAACGACAGGTATGTTGTC 300

Qy  360 ACATGTTATGAATGGTGATAAAAAGAGACTTCATTTTAATGGTGATGGCACAGGTATGGG 419
           |||||     |||||  || ||   | ||||||  ||| ||||||  ||||||| ||||
Db  301 TCAATGTCTGTCTCACATTACAGATACCAGCTTCATTGATTGGAGATGGGACAGGTCCTGG 360

Qy  420 ATTGCCTGAAAGTGAATTGATTATTTCGCTTTGAGGCATGTGAATATAGACGTCACTTTTTAAG 479
           |||||    ||||  |||      |  |     ||||     |||||||  ||||||
Db  361 TTCGGCCAATGCCAAATATTTTGTCTTTTGAATCTGACGAATATGAGGCGTCACTTCATGCC 420

Qy  480 TTATAAACCTGATTACGCGAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAA 539
           |||       || ||  ||||| |  ||||  ||   ||     |||||  |||||||||
```

```
Db  421 TTACCACCCAGAGATACTCTATTATCACCAACATTGACTTTGACCATCCAGATTATTCAC 480

Qy  540 AGATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATATAATGTTAAAAAGGTAT 599
            |||  ||| |||||||| |||||  ||||| | |  |        ||||
Db  481 AAGTTCCGAGGATGTTTTTAACGCCTTTAACGACTATGCAAACAAATCACCAAGGGTCT 540

Qy  600 TATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAGCAGAGTGTTCCAATTTATTA 659
            |||  | |  ||||||| ||  || |   |||| ||   ||   |    |||||||||
Db  541 TTTTGTCTATGGTGAAGATGCTGAATTGCCTAAGATTACGTCTGATGCACCAATTTATTA 600

Qy  660 TTATGGATTTAAAGATTCGG---ATGACATTTATGCTCAAAATATTCAAATTACCGGATAA 716
           |||||  |||  |  | | ||   ||||| ||| |
Db  601 TTATGGTTTTGAAGCTGAAGGCAATGACTTTGTAGCTAGTAGTGATCTTCTTCGTCAATAAC 660

Qy  717 AGGTACTGCTTTTGATGTGTATGGTGGATGGGATGTTTTATGATGATCACTTCCTGTCTCCCACA 776
            ||| |   ||||   ||   |   | |    |||
Db  661 TGGTTCAACCTTCACCGTTCATTTCCGTGGACAAAACTTGGGGCAATTCCACATTCCAAC 720

Qy  777 ATATGGTGACCATACAGTTTTAAATGCATTAGTCGTGAATTGCGATTAGTTATTTAGAGAA 836
           ||| |||     |||   ||  || ||  |||  |   |||||
Db  721 CTTTGGTCGTCACAATATCATGAATGCAGCCGTTATTGGTCTTCTTTACACAGCAGG 780

Qy  837 GCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTGTTAAACGTCGTTT 896
            | ||| | ||||   ||  ||| ||  ||| || ||||||  |||||  ||||||||||||
Db  781 ATTTGATTTGAACTTGGTGCGTGAGCACTTGAAAACATTGCCGGTGTTAAACGTCGTTT 840

Qy  897 CAATGAAACTACAATTGCAAATCAAGTTATTGTAGATGATTAATGCACACCATCCAAGAGA 956
           || |||||| | | ||| |||  |||  |||  |   |||  |  |||||||||||| ||||
Db  841 CACTGAGAAAAATTGTCAATGATACAGTGATTATTCGATGACTTTGCCCACCATCCAACAGA 900

Qy  957 AATTAGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGT 1016
           |||| | ||   |||||||  || | || || ||||| ||| |||| ||   ||| ||||
Db  901 AATTATTGCGACCTTGGATGCCTGTCGTCAGAAATACCCAAGCAAGAAGGAAATTGTAGCAGT 960

Qy 1017 ATTTCAACCACCACACTTTCTCTAGAACACAGCGCATTTTTAAATGAATTGCAGAAAGTTT 1076
            |||||||| |  | |||   | |||||| |  ||  ||||| | ||    |||
Db  961 CTTTCAACCGCATACCTTACAAGAACCATTGCCTTGTTGGACGACTTTGCCCATGCTTT 1020

Qy 1077 AAGTAAGACAGATCGTGTATTCGTTATGCTGAAATTTTTGGATCAATTAG---AGAAAATAC 1133
           ||   ||||||||  | |||    | ||  |||||||  ||| | |       ||||||||
Db 1021 AAACCAAGCAGATGCGTGTTTATCTTACGCGGACAAATTTATGGCTCCGGCTCGTGAAGTAGATCA 1080

Qy 1134 TGGGCGCATTAACGATCACAAGATTAAGTTAATTGATAAAT---TGAAGGTGCATCGTTAATTAA 1190
           |||  | ||||  | |  |||    |||  ||  |||    |   |||  |||||||
Db 1081 TGGTGACGTTAAGGTAGAAGACTAGCCAACAAAATCAACACAAAACAACCAAGTGATTAC 1140

Qy 1191 TGAAGATTCTATTAATGTTATTAGAACAATTTGATAATGCTGTTTATTTTATTATGGGTGC 1250
           |||| || | |  |  |||| | ||  ||| |||| ||||||||| |   ||| ||||| ||
Db 1141 TGTTGAAAATGTTTCTCCACTCCTAGACCATGACAATGCTGTTTACGGTCTTTTATGGGAGC 1200

Qy 1251 AGGTGATATTCAAAAATTACAAAATGCATAT 1281
           || |||||| ||||  ||||||  |   ||
Db 1201 AGGAGACATCCAAACCTATGAATACTCATTT 1231
```

```
RESULT 13
US-09-536-784-115
; Sequence 115, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
;  APPLICANT: Choi et. al.
;  TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;  NUMBER OF SEQUENCES: 452
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
```

```
;,..    OPERATING SYSTEM: MSDOS version 6.2
;,..    SOFTWARE: ASCII Text
;,..  CURRENT APPLICATION DATA:
;,..    APPLICATION NUMBER: US/09/536,784
;,..    FILING DATE: 30-Oct-1997
;,..    CLASSIFICATION: <Unknown>
;,..  PRIOR APPLICATION DATA:
;,..    APPLICATION NUMBER: 08/961,083
;,..    FILING DATE: OCT-30-1997
;,..  ATTORNEY/AGENT INFORMATION:
;,..    NAME: Michelle S. Marks
;,..    REGISTRATION NUMBER: 41,971
;,..    REFERENCE/DOCKET NUMBER: PB340P3
;,..  TELECOMMUNICATION INFORMATION:
;,..    TELEPHONE: (301) 309-8504
;,..    TELEFAX: (301) 309-8512
;,.  INFORMATION FOR SEQ ID NO: 115:
;,..    SEQUENCE CHARACTERISTICS:
;,..      LENGTH: 1267 base pairs
;,..      TYPE: nucleic acid
;,..      STRANDEDNESS: double
;,..      TOPOLOGY: linear
;,.    SEQUENCE DESCRIPTION: SEQ ID NO: 115:
US-09-536-784-115

Query Match     28.1%;    Score 370.2;  DB 4;  Length 1267;
Best Local Similarity  58.6%;    Pred. No. 5.9e-77;
Matches  721;  Conservative  0;  Mismatches 498;  Indels  12;  Gaps  4;

Qy      63 GCATGATTAGGACATGGAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAAGT 122
              ||| |||||||  |||  ||| | |||||| ||| ||||  || |||||||
Db       1 GCACCAGATGGGCACAAGGTTCAGGGATCAGATGTTGAAAAGTACTACTTTACCCAACG 60

Qy     123 TGCTCTTAGAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAAGA 182
              ||| | || |||  | ||||| |  ||||  |||||||| || || |||  ||||||
Db      61 CGGTCTTGAGCAGGCAGGAAATTACCATTCTCCTTTGATGAAAAAAATCTAGACGGTGA 120

Qy     183 TATGGTAGTTATACAAGGTAATGCATT---CGGCGAGTAGCCATGAAGAAATAGTACGTGC 239
              ||| |||  ||  || ||||| |||     ||| | || || || | ||| |||| |
Db     121 TATGGAAATTATCGCTGGAAATGCCTTTCGTCTTTGAATCTGACGAATAACGTGAAATTGCCTATGC 180

Qy     240 ACATCAATTGAAATTAGAGATGTTGTAAGTTATAAGATTTTTAGGACAGATTATTGATCA 299
              ||||||||   | |||   ||||  |  ||   |||||  |  ||||| |||||||
Db     181 GGACCAAAATGTATCAGCTACAAACGTTACCATGAGTTTCTAGGTAGCTTTATGCGTGA 240

Qy     300 ATATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATC 359
              ||||| | || |  |||  ||||| || |||| ||  ||| || | ||||||||||||
Db     241 CTTTGTTAGCATGGGAGTAGCAGGAGGAGCACATGGAAAACTTCAACGACAGGTATGTTGTC 300

Qy     360 ACATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGG 419
              ||| || || || ||| ||  | ||| |||  |  |||||| |||||| | || || ||
Db     301 TCATGTCTGTGTCCACATTACAGATACCAGCTTCTTCTTGATTGGAGATGGGACAGGTCGTG 360

Qy     420 ATTGCCTGAAGTGATTATTTCGCTTTTGAGGCGATGTGAATATAGACGTGCACTTTTTAAG 479
              ||||| ||  |  |||||||||||  ||  || ||| |  | || ||| | |||  |||
Db     361 TTCGGCCAATGCCAAATATTTTGTCTTTGAATCTGACGAATATGAGCGTGCACTTCATGCC 420

Qy     480 TTATAAACCTGATTACCGCCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAA 539
              |||  | | ||| |||  | | ||| ||  ||||| ||  || | || | ||||||| ||
Db     421 TTACCACCCAGAATAATCCTATTATCACCAACATTGACTTTGACCATCCAGATTATTTCAC 480

Qy     540 AGATATTAATAATGATGTTTTTGATGCATTCCAAGAAATGCACTATAATGTTAAAAAAGGTAT 599
              ||||||| | | |||| |||||||||||||  || || ||  ||  ||||||||||| ||
Db     481 AAGTCTCGAGATGTGTTTTAATGCCTTTAAGACTATGCCAAACAAATCACCAAGGGTCT 540

Qy     600 TATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTTATTA 659
              ||| || | ||| ||||| ||| ||| | | |||| || |||| || || ||||||||||
Db     541 TTTTGTCTATGTGAAGATGCTGAATTGCGTAAGATTACGTCTGATGCACCAATTTATTA 600

Qy     660 TTATGGATTTAAAGATTCGG---ATGACATTTATGCTCAAAATATTCAAATTACGGATAA 716
              |||| ||| | |||| ||     |||| ||| ||||| ||| || ||| | ||| || ||
Db     601 TTATGGTTTTGAAGCTGAAGGCAATGACTTTGTAGCTAGTAGTAGACTTTGTTTCGTTCAATAAC 660
```

;
RESULT 14
US-09-107-532A-3260
; Sequence 3260, Application US/09107532A
; Patent No. 6583275
;.. GENERAL INFORMATION:
;.. APPLICANT: Lynn A Doucette-Stamm and David Bush
;.. TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;..   ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;.. NUMBER OF SEQUENCES: 7310
;.. CORRESPONDENCE ADDRESS:
;..   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;..   STREET: 100 Beaver Street
;..   CITY: Waltham
;..   STATE: Massachusetts
;..   COUNTRY: USA
;..   ZIP: 02354
;.. COMPUTER READABLE FORM:
;..   MEDIUM TYPE: CD/ROM ISO9660
;..   COMPUTER: PC
;..   OPERATING SYSTEM: <Unknown>
;..   SOFTWARE: ASCII
;.. CURRENT APPLICATION DATA:
;..   APPLICATION NUMBER: US/09/107,532A
;..   FILING DATE: 30-Jun-1998
;.. PRIOR APPLICATION DATA:
;..   APPLICATION NUMBER: 60/085,598
;..   FILING DATE: 14 May 1998
;..   APPLICATION NUMBER: 60/051571
;..   FILING DATE: July 2, 1997
;.. ATTORNEY/AGENT INFORMATION:
;..   NAME: Ariniello, Pamela Deneke
;..   REGISTRATION NUMBER: 40,489

```
Qy     717 AGGTACTGCTTTGATGTGTATGTGGATGGTAGTTTATGATGATCACTTCCTGTCTCCACA 776
               ||||| |||||||| |||      |||||| ||||| | ||||| ||||| |||
Db     661 TGGTTCAACCTTCACCGGTTCATTTCCGTGTGACAAAACTTGGGCAATTCCACATTCCAAC 720

Qy     777 ATATGGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAA 836
               |   || ||    |||| ||  ||| |||| || ||| || || ||| |||  |||||
Db     721 CTTTGGTCGTCACAATATCATGAATGCGACAGCCGTTATTGGTCTTCTTTACACAGCAGG 780

Qy     837 GCTAGATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTTAAACGTCGTT 896
               |  ||  |||| ||  ||| |||| || || || ||||||| ||| ||||| || |||
Db     781 ATTTGATTTGAACTTGGTGCGTGAGCACTTGAAAAACATTTGCCCGGTGTTAAACGTCGTTT 840

Qy     897 CAATGAAACTACAATTGCAAATCAAGTTATTGTGAGATGATTATGCACACCATCCAAGAGA 956
               ||| ||  |  |||| ||||  | |  ||| ||| ||||| |||||  |||| ||| |
Db     841 CACTGAGAAAATTGTCAATGATACAGTGATTATCGATGACTTTGCCCACCACCAACAGA 900

Qy     957 AATTAGTGCTACAATTGAAACAGCACGAGAAGAAATATCCACATAAAGAAGTTGTTGCAGT 1016
               |||| |||  |||| ||| | ||  || |||  || |||||   ||| ||||| ||||||
Db     901 AATTATTGCGACCTTGGATGCGGCTCGTCAGAAATACCCAAGCAAGCAAGGAAATTGTAGCAGT 960

Qy    1017 ATTTCAACCACCACTTTCTTCTAGAACACACAGGCATTTTTAAATGAATTTGCGAGAAAGTTT 1076
               |||| |||| ||   |||    ||||||  |||| ||||  | || ||| || | |||||
Db     961 CTTTCAACCGCATACCTTTACAAGAACCATTGCCCTTGGTGTTGGACGACTTTGCCCATGCTTT 1020

Qy    1077 AAGTAAAGCAGATCGTGTGTATTCTTATGTGAAATTTTTGGATCAATTAG---AGAAAATAC 1133
               ||    || |  || |||| | ||| | |  || ||| ||| |  ||
Db    1021 AAACCAAGCAGATGCTGTTTATCCAGCGAAATTTATGGCTCGGCTCGTGAAGTAGATCA 1080

Qy    1134 TGGCGCATTAACGATACAGATAATTAATTGATAAAT---TGAAGGTGCATCGTGTTAATTAA 1190
               ||| |||  ||| ||| || |||| | |  || |||     ||| | ||||| |||| |||
Db    1081 TGGTGACGTTAAGGTAGGAAGACCTAGCCAACAATTCAACAAAAACACCAAGTGATTAC 1140

Qy    1191 TGAAGATTCTATTAAATGTATTAGAACAATTTGATGAAATGCTGTTTATTTTATTATGGGTGC 1250
               || |  |||  |||| ||| || | ||| || ||||  || || |||| ||||| |||||||
Db    1141 TGTTGAAAATGTTTCTCCACTCCTAGACCATGACCATGACAATGCTGTTTTACGTCTTTTATGGGAGC 1200

Qy    1251 AGTGATATTCAAAATTACAAAATGCATAT 1281
               || |  |||| ||||| || ||||| |||
Db    1201 AGGAGACATCCAAACCTATGAATACTCATTT 1231
```

```
                                                                    730 GATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATATGGTGACCAT 789
Qy                                                                      |||| ||| ||  |||  |||||   ||||||  ||||||||||
                                                                    757 GATGTCTCATCATGGGGATGAATTTGTTGGGCACTTCACTGTACCAGCATTCGGTAAACAC 816
Db

Qy                                                                  790 ACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTAGAGAAGAAGCTAGATGT 845
                                                                        ||  ||  |||||||||||||||||||||||||
Db                                                                  817 GATATCTTAAATGCTTTAGGTGTTATTGCGAGTAGCTTATGTCGAGAAACTGGATCT 872
```

RESULT 15
US-08-940-572-3/c
; Sequence 3, Application US/08940572
; Patent No. 6310193

US-08-940-572-3

Query Match   23.6%;   Score 310.4;   DB 4;   Length 1825;
Best Local Similarity   59.0%;   Pred. No. 5e-63;
Matches   607;   Conservative   0;   Mismatches 411;   Indels   10;   Gaps   4;

```
Qy    264 AAGTTATAATGATTTTTTAGGACAGATTATTGATCAATATACTTCAGTAGCTGTAACTGG 323
          |||      |  ||||| |   |||||   ||  |||||||||  ||||||||||||
Db   1797 ACGTTACCATGAGTTCTCAGGTAGCTTTATGCCTGACTTTGTTAGCATGGGAGTAGCAGG 1738

Qy    324 TGCACATGGTAAAACTTCTACAACAGAGTTATTATCACATGTTATGAATGGTGATAAAAA 383
          |||   ||  |||       |||  ||||| ||||||||   |||  ||
Db   1737 AGCACATGGAAAAAACTTCAACGACAGCAGGTATGTTGTCTCATGTCTGTCTCACATTACAGA 1678

Qy    384 GACTTCATTTTTAATTGGTGATGGCACAGGTATGGCAGGTATGGGATTGCCTGAAAGTGATTGATTATTTCGC 443
```

US-09-107-532A-3260

Query Match   26.9%;   Score 354.4;   DB 4;   Length 873;
Best Local Similarity   64.0%;   Pred. No. 2.4e-73;
Matches   535;   Conservative   0;   Mismatches 301;   Indels   0;   Gaps   0;

```
Qy     10 TATCATTTTGTCCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATCATGCATGAT 69
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     37 TATCATTTTGTTGGTATCAAAGGTTCAGGGATGAGCTCCTTGGCTCTTGTTTTACATGAA 96

Qy     70 TTAGGACATGAAGTTCAAGGATGCGGATCTGGAACTACGTATTACGAGAGTTGCTCTT 129
          ||||   ||||||  ||   |||||  ||  |||||  |||| |||||
Db     97 CAAGGTCTAAATGTCCAAGGATCAGACATCAGAAAAATATTTCTTTACACAAGAGATTTA 156

Qy    130 AGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAAGATATGGTA 189
          |||     ||||||||| |||||  ||  |||  |||||||||||||  |||
Db    157 GAAAAAGCAAATATCACGATTCTGCCATTTAACGCGAGATATAATGTAAAACCAGGGATGACG 216

Qy    190 GTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAGAAATAGTACGTGCACATCAATTG 249
          |||| ||||||||||||||||  ||  |||  || ||||||||||
Db    217 ATCATTGCAGGAAATGCATTTCCTGATTCACATGAACATGAAGAAATCCAGCGAGCAAAAGAATTA 276

Qy    250 AAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAATAATACTTCA 309
          ||| |  ||||| |||||||||||  |||  ||||||||||| |||||||||||||
Db    277 GGTTTGGAAGTGATTCGCTATCACGACTTTATTGGTCATTTATCCAAAATTATACGAGT 336

Qy    310 GTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCACATGTTATG 369
          |||||      ||||  |||||||
Db    337 ATCGCTGTTAACAGGTCTCACGGAAAAAACAGTACGACTGCTGCTCTCTCATGTATTG 396

Qy    370 AATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGATTGCCTGAA 429
          |||||   |||||   |  || |||  ||||  ||||||||||||||
Db    397 AGTGGTGTCCGTCCAACAGTTATCTAATTGCTAATTGGAGATGGAACAGGACATGGTGACATCCGGCAA 456

Qy    430 AGTGATTATTTCGCTTTTGAGGCATGTGAATATGACGTCACTTTTTAAGTTATAAACCT 489
          ||  |||  |||  ||| |  ||||||  ||||  | |  |||  ||||||
Db    457 GCCGGAATTCTTTCTTATTGCCGGATTTGAGGCCTGTGAATATCGCCGTCATTTCCTTGCCTATTCACCA 516

Qy    490 GATTACCGCAATTATGACAAAATATTGATTTCGATCATCCTGATTAATTTAAAGATATTAAT 549
          |||||   |||||||    |||||||   |||||||| ||||||||| || |  |||
Db    517 GACTATGCAATCATGACGACCAATATCGATTTTTGATCATCCGGATTATTATTATRCAAGTATCGAC 576

Qy    550 GATGTTTTGATGCATTCCAAGAAAATGGCACATAAATGTTAAAAAAGGTTATAATTGCTTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    577 GATGTATTTACAGCTTTCCAAACAATGGCTGGACAATGGGAAAAAAAGCTATTTTTGCGTAT 636

Qy    610 GGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTTATTATTGGATTT 669
          ||  |||||   ||||||   |||  |||   |||||     ||||   ||| |  ||||
Db    637 GGTGACGATGCTTATTACGACGAATTAAAGCAAATGTGCCGATTTATTTATTATTATGGTGTC 696

Qy    670 AAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAGTACTGCTTTT 729
          |||  |  |||  |||||  |||  ||  ||||||||||||||||||||
Db    697 ACAGAAAACGATGATATCCAACGACGAAACATCGAACGACCACAACTTCTGGTTCTGTTCGTTCTTTT 756
```

Db   1677 TACCAGCTTCTTGATTGGAGATGGGACAGGTCGTGGTTCGGCCAATGCCAAATATTTTGT 1618
              |||| ||||||||||| |||||||||| ||||||| ||||||||||||| || |||||||
Qy   444 TTTTGAGGCATGTGAATATAGACGTCACTTTTAAGTTATAAACCTGATTACGCCAATTAT 503
              ||||| ||||| |||||| ||| |||||||||| |||||||| || |||| ||||||||
Db   1617 CTTTGAATCTGACGAATATGAGCGTCACTTCACTTCATGCCTTACCCAGAATACTTCTATTAT 1558

Qy   504 GACAAATATTGATTTCGATCATCCTGATTATTTAAAGATATTAAATGATGTTTTGATGC 563
              ||||||||| |||||| || ||| ||||||| |||||||||| |||||||||||||||||
Db   1557 CACCAACATTGACTTTGACCATCCAGATTATTTCACAAGTCTCGAGGATGTTTCAATGC 1498

Qy   564 ATTCCAAGAAATGGCACATAATGTTAAAAAAGTATTATTGCTTGGGGTGATGATGAACA 623
              |||| || | |||| || |||| ||| ||||| || || || |||||||||||| |
Db   1497 CTTTAACGACTATGCCAAACAAATTACGACGGGTCTTTTTGTCTATGCTGAAGATGCTGA 1438

Qy   624 TCTACGTAAAATTGAAGCAGATGTTCCAATTTATTATTATGATTAAAGATTCGG---A 680
              ||||| ||||||||||||| |||| ||||||||||| |||||||| |||||| ||
Db   1437 ATTGCGTAAGATTACCTCTGATGCACCAATTTATTATTATGTTTGAAGCTGAAGGCAA 1378

Qy   681 TGACATTTATGCTCAAAATATTCAAATTACGGATAAAAGGTACTGCTTTGATGTGTATGT 740
              ||||| |||||| |||| | || ||||| ||||| |||| || ||||||| ||||| ||
Db   1377 TGACTTTGTAGCTAGTGATCTTCTTCGTTCAACAACTGGTTCAACCTTCACCGTTCATTT 1318

Qy   741 GGATGGCTGAGTTTTTATGATCACTTCCGTGTCTCCACAATATGTGACCATACAGTTTTAAA 800
              || |||| ||| |||||| ||| |||| || || ||| || || ||| ||||||||
Db   1317 CCGTGGACAAAACTTGGGGCAATTCCACATTCCAACCTTGGTTCGTCACAATATCATGAA 1258

Qy   801 TGCATTAGCTGTGTAATTGCGATTAGTTATTTGAGAAGCTAGAGTTTACAAATATTAAAGA 860
              || ||| |||| |||| | ||| ||| | | |||||| |||| ||| | || |||||
Db   1257 TGCCACAGCCGTTATTGGTCTTCTTCTTTACACAGCAGGATTTGATTTGAACTTGGTGCGTGA 1198

Qy   861 AGCATTAGAAACGTTTGGTGGTGTAAACGTTGGTTGCTGTTCAATGAAACTACAATTGCAAATCA 920
              ||| || | |||| || |||||||||||||| ||| ||||| ||||| ||||||||||| ||
Db   1197 GCACTTGAAAACATTTGCCGGGTGTTAAACGTCGTTCACTGAGAAAATTGTCAATGATAC 1138

Qy   921 AGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGCTACAATTGAAACAGC 980
              |||||||||| |||||||||||| ||| ||||| |||||| || || ||| || |||||
Db   1137 AGTGATTATTGATGACTTTGCCCATCATCCAACAGAAATTATTGCGACCTTGGATGCGGC 1078

Qy   981 ACGAAAGAAATATCCACATAAAGAAGTTGTTGCAG-TATTTCAACCACCACACTTTCTCTA 1039
              ||| || | |||| | ||| |||| |||||| |||||||| || |||| ||||
Db   1077 TCGTCAGAAATACCCCAAGGAAATTGTAGCAGTTGTAGCAGTTCTTCCACCGCCATACCTTTACAA 1018

Qy   1040 GAACACAGGCATTTTAAATGATTGCAGAAAGTTAAGTAAAGCAGATTCGTGTATTCT 1099
              ||||||| | || ||| | || |||| ||| || || ||| |||| || |||||
Db   1017 GAACCATTGCCCTGTTGGACGACTTTGCCCATGCTTTAACCAAGCAGATGCTGTTTATC 958

Qy   1100 TATGTGAAATTTTGGATCAATTAG---AGAAAATACTGGCGCATTAACGATACAAGATT 1156
              ||||| || ||| |||| |||||| |||||||| ||| |||| |||| || ||| ||
Db   957 TAGCGCAAATTTATGGCTCGGCTCGTGAAGTGATCATCGTGACGTTAAGGTAGAAGACC 898

Qy   1157 TAATTGATAAAAT---TGAAGGTGCATCGTTAATTAAATGAAGATTCTATTAATGTATTAG 1213
              |||| || ||||| || || ||| | ||| || |||||| | |||| |||| |||||
Db   897 TAGCCAATAAAATCAACAACAAAAACACCAAGTGATTACTTACTTGTTGAAAATGTTTCTCCACTCC 838

Qy   1214 AACAATTTGATAATGCTGTTATTTTATTTATGGGTGCAGGTGATGATTCAAAAATTACAAA 1273
              ||||||||| || ||| |||| || |||||||||||||||| |||| || ||||||||||
Db   837 TAGACCATGCAAATGCTGTTTTACGTCTTTATGGGAGCAGGAGACATCCAAACCTATGAAT 778

Qy   1274 ATGCATAT 1281
              |||||||
Db   777 ACTCATTT 770

Search completed: June 6, 2004, 02:47:06
Job time : 142 secs

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 02:47:10 ; Search time 5352 Seconds
                                          (without alignments)
                                          10673.789 Million cell updates/sec

Title: US-10-712-713-1
Perfect score: 1318
Sequence: 1 atgacacactatcattttgt........gaaaaatgcgttttaagctt 1318

Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167151695 residues
Word size :   50

Total number of hits satisfying chosen parameters: 17

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Result No.		Score	% Match	Query Length	DB	ID	Description
1		1314	99.7	1335	6	AX742069	AX742069 Sequence
2		1314	99.7	2424	6	AR354274	AR354274 Sequence
3		1311	99.5	1332	6	AX617936	AX617936 Sequence
4		1263	95.8	1335	6	AX742071	AX742071 Sequence
5	c	1212	92.0	301550	1	AP003134	AP003134 Staphyloc
6	c	1212	92.0	342600	1	AP003363	AP003363 Staphyloc
7	c	1161	88.1	272850	1	AP004828	AP004828 Staphyloc
8		995	75.5	1314	6	AF034076	AF034076 Staphyloc
9		995	75.5	1314	6	AX191741	AX191741 Sequence
10		842	63.9	1351	6	E35613	E35613 MurC. 6/200
11		842	63.9	1351	6	BD178030	BD178030 MurC. 4/2
12		434	32.9	660	6	E35614	E35614 MurC. 4/2
13		434	32.9	660	6	BD178031	BD178031 MurC. 4/2
14	c	275	20.9	619	6	AR194630	AR194630 Sequence
15	c	90	6.8	888	6	AR149358	AR149358 Sequence
16	c	90	6.8	888	6	AR405068	AR405068 Sequence
17	c	90	6.8	888	6	AR411614	AR411614 Sequence

Database :   GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                10: gb_ro:*
                11: gb_sts:*
                12: gb_sy:*
                13: gb_uni:*
                14: gb_vi:*
                15: em_ba:*
                16: em_fun:*
                17: em_hum:*
                18: em_in:*
                19: em_mu:*
                20: em_om:*
                21: em_or:*
                22: em_ov:*
                23: em_pat:*
                24: em_ph:*
                25: em_pl:*
                26: em_ro:*
                27: em_sts:*
                28: em_un:*
                29: em_vi:*
                30: em_htg_hum:*
                31: em_htg_inv:*
                32: em_htg_other:*
                33: em_htg_mus:*
                34: em_htg_pln:*
                35: em_htg_rod:*
                36: em_htg_mam:*
                37: em_htg_vrt:*
                38: em_sy:*
                39: em_htgo_hum:*
                40: em_htgo_mus:*
                41: em_htgo_other:*

ALIGNMENTS

RESULT 1
AX742069
LOCUS   AX742069   1335 bp   DNA   linear   PAT 10-MAY-2003
DEFINITION   Sequence 27 from Patent WO03025007.
ACCESSION   AX742069
VERSION   AX742069.1 GI:3052456
KEYWORDS   .
SOURCE   Staphylococcus aureus
  ORGANISM   Staphylococcus aureus
              Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Edwards,A., Dharamsi,A., Vedadi,M., Alam,M.Z., Awrey,D.,
              Beattie,B., Domagala,M., Houston,S., Kanagarajah,D., Nethery,K.,
              Ng,I., Mansoury,K., Mcdonald,M.L., Pinder,B., Viola,C. and
              Wrezel,O.
  TITLE   Novel purified polypeptides involved in membrane biosynthesis
  JOURNAL   Patent: WO 03025007-A 27 27-MAR-2003;
              Affinium Pharmaceuticals, Ind. (CA)
FEATURES   Location/Qualifiers
     source   1..1335
                     /organism="Staphylococcus aureus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1280"
ORIGIN

Query Match   99.7%; Score 1314; DB 6; Length 1335;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1314; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 81

Qy   61 ATGCATGATTTAGGACATGAAGTTCAAGGATCCGGATATTGAGAACTACGTATTTACAGAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATTTAGGACATGAAGTTCAAGGATCCGGATATTGAGAACTACGTATTTACAGAA 141

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 201

Qy   181 GATATCGTAGTTATACAAGTAATGCATTCGCCAGTAGCCATGAAGAAATAGTACGTGCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Pred. No. is the number of results predicted by chance to have a

```
RESULT 2
AR354274
LOCUS       AR354274                2424 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 392 from patent US 6593114.
ACCESSION   AR354274
VERSION     AR354274.1  GI:33760358
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2424)
  AUTHORS   Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6593114-A 392 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1..2424
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN

Query Match     99.7%;  Score 1314;  DB 6;  Length 2424;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1314;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGACACHATCATCATTTGTCCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 ATGACACHATCATCATTTGTCCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 686

Qy     61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 746

Qy    121 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTGATGCTAATAACATAAAAGAA 806

Qy    181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGTATAATGATTTTTAGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGTATAATGATTTTTAGGA 866

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTGTAAGTGTAAACTTCTCACACAGGTTATTATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 CATCAATTGAAATTAGATGTTGTAAGTTGTAAGTGTAAACTTCTCACACAGGTTATTATCA 926

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGTAAAAAGACTTCATTTTAATGGTGATGCGCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 TATACTTCAGTAGCTGTAACTGGTGCACATGTAAAAAGACTTCATTTTAATGGTGATGCGCA 986

Qy    361 CATGTTATGAAATGGTGATAAAAGAAGATTGTTGTTTGCAGGTATGGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 CATGTTATGAAATGGTGATAAAAGAAGATTGTTGTTTGCAGGTATGGGA 1046

Qy    421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGGAGGCATGGCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1047 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGGAGGCATGGCACTTTTTAAGT 1106

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1107 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 1166

Qy    541 GATATTAATGATGTTTTGATGATGAACATCACGTAATAATTACGGCACATAATGTGATATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1167 GATATTAATGATGTTTTGATGATGAACATCACGTAATAATTACGGCACATAATGTGATATT 1226

Qy    601 AATTGCTTGGGGTGATGATGAACAGATGTTCCAATTACGGATAAAGGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1227 AATTGCTTGGGGTGATGATGAACAGATGTTCCAATTACGGATAAAGGT 1286

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAAATACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    202 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAATAGTACGTGCA 261
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGACAGATTGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGACAGATTGATCAA 321

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCTACAACAGGTTTATTATCA 381

Qy    361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGGTGATGGCACAGGTATGGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGGTGATGGCACAGGTATGGGA 441

Qy    421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGGAGTGAATATAGACGTCACTTTTAAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGGAGTGAATATAGACGTCACTTTTAAGT 501

Qy    481 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 561

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCCACATATAATGTTAAAAAAGGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCCACATATAATGTTAAAAAAGGTATT 621

Qy    601 AATTGCTTGGGGTGATGATGAACATCACGCTAAAATTGAAGCAGATGTTCCAATTTATTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 AATTGCTTGGGGTGATGATGAACATCACGCTAAAATTGAAGCAGATGTTCCAATTTATTAT 681

Qy    661 TATGGATTTAAAGATCTGGATGTGTATGTCGGATGTGAGTTTATGATGTCTGAGTTATAGCT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGATTTAAAGATCTGGATGTGTATGTCGGATGTGAGTTTATGATGTCTGAGTTATAGCT 741

Qy    721 ACTGCTTTTGATGTGTATGTCGGATGTGAGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTTGATGTGTATGTCGGATGTGAGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTGAATTGCATTAGCAGAGAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTAAATGCATTAGCTGTGAATTGCATTAGCAGAGAGCTA 861

Qy    841 GATGTTACAAAATATAAAAGAAGCATTAGAAACGTTGGTGTGTTGTAAACGTCGTTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAAATATAAAAGAAGCATTAGAAACGTTGGTGTGTTGTAAACGTCGTTTCAAT 921

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGAATGATGATGATCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAATTGCAAATCAAGTTATTGTAGAATGATGATGATCAAGAGAAATT 981

Qy    961 AGTGCTACAACACTTTCTCTAGAAACACAGCACGAAAGAAATATCCACATAAAGAAGATTCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTGCTACAACACTTTCTCTAGAAACACAGCACGAAAGAAATATCCACATAAAGAAGATTCT 1041

Qy   1021 CAACCACACACTTTCTCTAGAACACAGCAGCACTTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCTAGAACACAGCAGCACTTTTAAATGAATTTGCAGAAAGTTTAAGT 1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGTAGAATTAGAGAAAATTTTGATCAATTAGAGAAAATT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGTAGAATTAGAGAAAATTTTGATCAATTAGAGAAAATT 1161

Qy   1141 TTAACGATACAAGATTAATTGAAATGTGCATCCGTTAATTAATGGGTCAGGTCAGGAGATTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TTAACGATACAAGATTAATTGAAATGTGCATCCGTTAATTAATGGGTCAGGTCAGGAGATTCT 1221

Qy   1201 ATTAATGATTATTAGAACAATTTGATAATGCTGTGTATTTTATTTAATGGGTGCAGGTGATATT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 ATTAATGATTATTAGAACAATTTGATAATGCTGTGTATTTTATTTAATGGGTGCAGGTGATATT 1281

Qy   1261 CAAAAATTACAAACATCATATTTAGATAATAGCATGAAAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 CAAAAATTACAAACATCATATTTAGATAATAGCATGAAAAAATGCGTTTTAA 1335
```

us-10-712-713-1.oligo.rge

```
Db   1287 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAAGGT 1346
Qy    721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 1406

Qy    781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCAGATTAGTATTTAGAGAAGCTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCAGATTAGTATTTAGAGAAGCTA 1466

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTAAACGTCGTTTCAAT 1526

Qy    901 GAAACTACAATTGCAAATCAAGTCAAATCAAGTTATTGTAGATGATAATGCACACCATCCAAGAGAAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 GAAACTACAATTGCAAATCAAGTCAAATCAAGTTATTGTAGATGATAATGCACACCATCCAAGAGAAATT 1586

Qy    961 AGTGCTACAAGTACAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 AGTGCTACAAGTACAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1646

Qy   1021 CAACCACACACTTTCTCTAGAACACAGGCCATTTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 CAACCACACACTTTCTCTAGAACACAGGCCATTTTTAAATGAATTTGCAGAAAGTTTAAGT 1706

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCGCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCGCA 1766

Qy   1141 TTAACGATACAAGATTTAATTGATAAAATTGATGAAGTGCATCGTTAATTAATGAAGATTCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1767 TTAACGATACAAGATTTAATTGATAAAATTGATGAAGTGCATCGTTAATTAATGAAGATTCT 1826

Qy   1201 ATTAATGTATTAGAACAATTGATPAATGCTGTTATTTATTTATTATGGTGCAGGTGATATT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 ATTAATGTATTAGAACAATTGATAATGCTGTTATTTATTTATTATGGTGCAGGTGATATT 1886

Qy   1261 CAAAAATTACAAAGTGCATATTTAGATAAATTAGGCATGAAAAAAATGCGTTTTAA 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1887 CAAAAATTACAAAGTGCATATTTAGATAAATTAGGCATGAAAAAAATGCGTTTTAA 1940
```

RESULT 3

AX617936

LOCUS	AX617936 1332 bp DNA linear PAT 20-FEB-2003
DEFINITION	Sequence 899 from Patent WO02094868.
ACCESSION	AX617936
VERSION	AX617936.1 GI:28448287
KEYWORDS	.
SOURCE	Staphylococcus aureus
ORGANISM	Staphylococcus aureus
	Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE	1
AUTHORS	Masignani,V.C., Mora,M.C. and Scarselli,M.C.
TITLE	Staphylococcus aureus proteins and nucleic acids
JOURNAL	Patent: WO 02094868-A 899 28-NOV-2002;
	Chiron Spa (IT)
FEATURES	Location/Qualifiers
source	1..1332
	/organism="Staphylococcus aureus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:1280"
ORIGIN	

```
Query Match      99.5%;  Score 1311;  DB 6;  Length 1332;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1311;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGACAGACACTATCATTTTGTCGGAATTAAAGGTTCTGCCATGAGTTCATTAGACACAATC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACAGACACTATCATTTTGTCGGAATTAAAGGTTCTGCCATGAGTTCATTAGACACAATC 81
```

```
Db     61 ATGGCATGATTTAGGACATGAAATTGAGAACTACGGATTATTACAGAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGGCATGATTTAGGACATGAAATTGAGAACTACGGATTATTACAGAA 141

Qy    121 GTTGCTTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GTTGCTTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 201

Qy    181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCAGTAGCCATGAAGAAGAAATAGTACGTGCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCAGTAGCCATGAAGAAGAAATAGTACGTGCA 261

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 321

Qy    301 TATACTTCAGTAGCTGTAACTGGTGCACATGTAAAACTTCACAACAGGTTTATTATCA 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 TATACTTCAGTAGCTGTAACTGGTGCACATGTAAAACTTCACAACAGGTTTATTATCA 381

Qy    361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTGCACTTTTTAAGT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTGCACTTTTTAAGT 501

Qy    481 TATAAACCTGATTACGCAATTATGACAAAATATGATTCGATCATCCTGATATTTTAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 TATAAACCTGATTACGCAATTATGACAAAATATGATTCGATCATCCTGATATTTTAAA 561

Qy    541 GATATTAATGATGTTTGATGCATTCCAAGAAATGGCACATATGTTAAAAAAGGTATT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATATGTTAAAAAAGGTATT 621

Qy    601 ATTGCTTGGGGTGATGATGAACATCACGTAAAATGAAGCAGATGTTCCAATTTATTAT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ATTGCTTGGGGTGATGATGAACATCACGTAAAATGAAGCAGATGTTCCAATTTATTAT 681

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAAGGT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGATTTAAAGATTCGGATGACATTTATGCTCAAAATATTCAAATTACGGATAAAAGGT 741

Qy    721 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy    781 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCAGATTAGTATTTAGAGAAGCTA 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTTAAATGCATTAGCTGTAATTGCAGATTAGTATTTAGAGAAGCTA 861

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTGTTAAACGTCGTTTCAAT 921

Qy    901 GAAACTACAATTGCAAATCAAGTCAAATCAAGTTATTGTAGATGATTATGTGATATGCACCATCCAAGAGAAATT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAATTGCAAATCAAGTCAAATCAAGTTATTGTAGATGATTATGTGATATGCACCATCCAAGAGAAATT 981

Qy    961 AGTGCTACAATTGAAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTGCTACAATTGAAAACAGCACGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACACACTTTCTCTAGAACACAGGTTCTGCGCATGAGTTCATTAGACACAAATC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCTAGAACACAGGTTCTGCGCATGAGTTCATTAGACACAAATC 1101

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCCCA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAAATACTGGCCCA 1161

Qy   1141 TTAACGATACAAGATTTAATTGATAAAATTGATGAAGTGCATCGTTAATTAATGAAGATTCT 1200
```

```
Db   481  TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTAAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   502  TATAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTAAA  561

Db   541  GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAAGGTATT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   562  GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAAGGTATT  621

Db   601  AATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGAGTTCCAATTTATTAT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   622  AATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  681

Db   661  TATGGATTTAAAGATTCGGATGACAGATTTATGCTCAAATATTCAAATACGGATAAAGGT  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   682  TATGGATTTAAAGATTCGGATGACAGATTTATGCTCAAATATTCAAATACGGATAAAGGT  741

Db   721  ACTGCTTTTGATGGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   742  ACTGCTTTTGATGGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  801

Db   781  GGTGACCATACAGTTTAAAATGCATTAGCTAGCTGTAATTGCGATTAGTAGTTATTAGAGAAGCTA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   802  GGTGACCATACAGTTTAAAATGCATTAGCTAGCTGTAATTGCGATTAGTAGTTATTAGAGAAGCTA  861

Db   841  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGGTGTGTTAAACGTCGTTTCAAT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   862  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGGTGTGTTAAACGTCGTTTCAAT  921

Db   901  GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTGTGCACACCATCCAAGAGAAATT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   922  GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTGTGCACACCATCCAAGAGAAATT  981

Db   961  AGTGCTACAAATTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   982  AGTGCTACAAATTGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT  1041

Db   1021 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGAAGAAAGTTTAAGT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1042 CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGAAGAAAGTTTAAGT  1101

Db   1081 AAAGCAGATCGTGTATCTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1102 AAAGCAGATCGTGTATTCTTATGTGAAATTTTGGATCAATTAGAGAAAATACTGGCGCA  1161

Db   1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATTAATGAAGATTCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1162 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1221

Db   1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTGGGTGCAGTGGATATT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1222 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTTATTTGGGTGCAGTGGATATT  1281

Db   1261 CAAAAATTACAAAATGCATATTTAGATAAAAATTAGGCATGAAAAATGCGTTTAA  1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1282 CAAAAATTACAAAATGCATATTTAGATAAAAATTAGGCATGAAAAATGCGTTTAA  1335
```

```
RESULT 4
AX742071
LOCUS        AX742071              1335 bp    DNA       linear   PAT 10-MAY-2003
DEFINITION   Sequence 29 from Patent WO03025007.
ACCESSION    AX742071
VERSION      AX742071.1  GI:30524567
KEYWORDS     .
SOURCE       Staphylococcus aureus
  ORGANISM   Staphylococcus aureus
             Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE    1
  AUTHORS    Edwards,A., Dharamsi,A., Vedadi,M., Alam,M.Z., Awrey,D.,
             Beattie,B., Domagala,M., Houston,S., Kanagarajah,D., Nethery,K.,
             Ng,I., Mansoury,K., Mcdonald,M.L., Pinder,B., Viola,C. and
             Wrzel,O.
  TITLE      Novel purified polypeptides involved in membrane biosynthesis
  JOURNAL    Patent: WO 03025007-A 29 27-MAR-2003;
             Affinium Pharmaceuticals, Inc. (CA)
FEATURES             Location/Qualifiers
     source          1..1335
                     /organism="Staphylococcus aureus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1280"
ORIGIN

Query Match     95.8%;  Score 1263;  DB 6;  Length 1335;
Best Local Similarity   99.9%;
Matches 1313;    Conservative    0;   Mismatches    0;   Indels    1;   Gaps    0;

Qy      1  ATGACACACTATCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22  ATGACACACTATCATCCATTTGTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC  81

Qy     61  ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82  ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA  141

Qy    121  GTTGCTCTTAGAGAAATAAGGGGATAAAATAAATATTACCATTGATGCTAATAACATAAAGAA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142  GTTGCTCTTAGAGAAATAAGGGGATAAAAAATATTACCATTGATGCTAATAACATAAAAGAA  201

Qy    181  GATATGGTAGTTATACAAGGTAATGCATTCCGCGAGTAATGCATTCCGCGAGTAGCGTGCA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202  GATATGGTAGTTATACAAGGTAATGCATTCCGCGAGTAATGCATTCCGCGAGTAGCGTGCA  261

Qy    241  CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262  CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTTAGGACAGATTATTGATCAA  321

Qy    301  TATACTTCAGTCAGTCGTCAACTGGTGCACATGGTAAAACTTCTACACACAGGTTTATTATCA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322  TATACTTCAGTCAGTCGTCAACTGGTGCACATGGTAAAACTTCTACACACAGGTTTATTATCA  381

Qy    361  CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382  CATGTTATGAATGGTGATAAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA  441

Qy    421  TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAGT  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442  TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTTAAGT  501
```

```
RESULT 5
AP003134/c
LOCUS        AP003134             301550 bp    DNA       linear   BCT 24-APR-2003
DEFINITION   Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
             genome, section 6/10.
ACCESSION    AP003134 BA000018
VERSION      AP003134.2  GI:14349226
KEYWORDS     .
SOURCE       Staphylococcus aureus subsp. aureus N315
  ORGANISM   Staphylococcus aureus subsp. aureus N315
             Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE    1
  AUTHORS    Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
             Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
             Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
             Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
```

```
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
   MEDLINE  21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 301550)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES             Location/Qualifiers
     source          1..301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /mol_type="genomic DNA"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene            complement(249..947)
                     /gene="SA1293"
     CDS             complement(249..947)
                     /gene="SA1293"
                     /note="ORFID:SA1293"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42553.1"
                     /db_xref="GI:13701259"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIMLGIILTAICSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVRATAVAV
                     AELVRIILIARSSD"
     gene            complement(1001..1588)
                     /gene="SA1294"
     CDS             complement(1001..1588)
                     /gene="SA1294"
                     /note="ORFID:SA1294"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42554.1"
                     /db_xref="GI:13701260"
                     /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                     HFKIFVPDSPTAILYIVISISLILIQKQNSIIEALAFVTLFYGIWAVIMNLFIIEQ
                     GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                     DFITHVWQIGVLSCCLSVFGLLLVIELNKLLKCK"
     gene            complement(1578..2153)
                     /gene="SA1295"
     CDS             complement(1578..2153)
                     /gene="SA1295"
                     /note="ORFID:SA1295"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42555.1"
                     /db_xref="GI:13701261"
                     /translation="MSETLNQIKESFIEYLLFQYRFKSRIAVWVLNYIKVNEAKLANI
                     HFVDTKINHHETELEIAEVGSHASAIQFTKRNIKLMNTNEIFDYIANHNCAFDIQIHFA
                     NVSKREQRLDDLIVVAQLTESPSYQYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                     EKQRFYQLTQILNTLKLVNKHQFEDLLADDD"
     gene            complement(2167..3411)
                     /gene="SA1296"
     CDS             complement(2167..3411)
                     /gene="SA1296"
                     /note="ORFID:SA1296"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42556.1"
                     /db_xref="GI:13701262"
                     /translation="MEDIYKLIDDINLQKLENLDSRVNEAITTDNDDALFIGETLYN
                     FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                     ADLYQQINMMEVAIDKLQEALELEPNDPIIHFALAEMLYYDGQYLRATSEYETVLETG
                     EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISYDKNDITQE
                     AIKIMTTLLSKDPYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTTGC
                     LEIEHGDANRGVLLKQALEVDNAYQEPLLLISDLYRNEEDYEAIIELLITYDEEDLD
                     PTFMWHLAFAYGQEERDKEAQHFFELAYPTMKTNIDFMSDYFYLIEIGQKEQAITIL
                     NQLLELEPSNENWHDELSRIQS"
     gene            complement(3418..4716)
                     /gene="aroA"
     CDS             complement(3418..4716)
                     /gene="aroA"
                     /note="ORFID:SA1297"
                     /codon_start=1
                     /transl_table=11
                     /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                     /protein_id="BAB42557.1"
                     /db_xref="GI:13701263"
                     /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
                     PLLGDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAPKTPHQVLYTGNSGTTRLLA
                     GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLIIKPSVIKGIN
                     YQMEVASAQVKSAILFALSFSNDTTVIKELDVSRNHTETMFRHFNIPIEAERLSTTT
                     PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                     GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                     IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                     MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene            complement(4726..5790)
                     /gene="aroB"
     CDS             complement(4726..5790)
                     /gene="aroB"
                     /note="ORFID:SA1298"
                     /codon_start=1
                     /transl_table=11
                     /product="3-dehydroquinate synthase"
                     /protein_id="BAB42558.1"
                     /db_xref="GI:13701264"
                     /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                     YFADKFDDILSYENVHKVIIPAGEKTKTFEQYQETLEYILSHVTRNTAIIAVGGGAT
                     GDFAGPVAATLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                     LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEOHFKDREILQSLNGMDKYIAKGI
                     ETKLDIVVADEKEQGVRKFLNLGHTFGHAVEYNHKIAHGHAVMIGIIYQFIVANILFN
                     SNHDIQHYINYLTKLGVPLETTDIDPETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
                     QHIDQTTLQHACEQLKTYFK"
     gene            complement(5816..6982)
                     /gene="aroC"
     CDS             complement(5816..6982)
                     /gene="aroC"
                     /note="ORFID:SA1299"
                     /codon_start=1
                     /transl_table=11
                     /product="chorismate synthase"
                     /protein_id="BAB42559.1"
                     /db_xref="GI:13701265"
                     /translation="MRYLTSGESHGPQLTVIVEGIPANLEIKVEDINKEMFKRQGYVG
                     RGRRMQIBKDTVEIVSGVRNGYTLGSPITMVVTNDDFTHWRKIMGAAPISEEERENMK
                     RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETAARVAVGALCKVLLQQLDIDIY
                     SRVVEIGGIKDKDFYDSETFKANLDRNDVRIDDSIAQAMRDKLNEGDSIGGVV
                     QVVVENMPVGVGSVYHYDRKLDGKIAQGVVSINAFKGVSFGEGFKAABKPGSEIQDEI
                     LYNSEIGYTRGSNHLGGLEGGMSNGMPIIVNGVMKPIPTTYKPLNSVDINTKEDFKAT
                     IERSDSCAVPAASIVCEHVVAFEIAKALLBEFQSNHIEQLQOQIADRRQLNVEF"
     gene            complement(7368..7568)
                     /gene="SA1300"
     CDS             complement(7368..7568)
                     /gene="SA1300"
                     /note="ORFID:SA1300"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
```

```
                /protein_id="BAB42560.1"
                /db_xref="GI:13701266"
                /translation="MKNIISIILGILMFLKLMELLYGAIFLDKPLNPITKIFILTLI
                YIFYVLVKELIIFLKSKYNKSA"
gene            complement(7597..7722)
                /gene="SAS045"
CDS             complement(7597..7722)
                /gene="SAS045"
                /note="ORFID:SAS045"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB42561.1"
                /db_xref="GI:13701267"
                /translation="MWLVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"
gene            complement(7777..8226)
                /gene="ndk"
CDS             complement(7777..8226)
                /gene="ndk"
                /note="ORFID:SA1301"
                /codon_start=1
                /transl_table=11
                /product="nucleoside diphosphate kinase"
                /protein_id="BAB42562.1"
                /db_xref="GI:13701268"
                /translation="MERTFLMIKPDAVORNLIGEVISRIERKGLKLVGGKLMQVPMEL
                AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGDAVNVSRHIIGSTNPSEASPGSIR
                GDLGLTVGRNIIHGSDSLESAREINLWFNENEITSYASPRDAWLYE"
gene            complement(8318..9277)
                /gene="gerCC"
CDS             complement(8318..9277)
                /gene="gerCC"
                /note="ORFID:SA1302"
                /codon_start=1
                /transl_table=11
                /product="heptaprenyl diphosphate syntase component II"
                /protein_id="BAB42563.1"
                /db_xref="GI:13701269"
                /translation="MAKLNMNNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
                PAFVLLSSQFGKDEQTSEQTYQVAVALEIHMATLVHDDVIDKSDKRRGKLTISKKWD
```

```
Query Match        92.0%;   Score 1212;   DB 1;   Length 301550;
Best Local Similarity  99.8%;  Pred. No. 0;
Matches 1312;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      1 ATGACACACTATCATTTTGTTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 291118 ATGACACACTATCATTTTGTTCGGAATTAAAGGTTCTCGGCATGAGTTCATTAGCACAAATC 291059

Qy     61 ATGGCATGATTAGGACAGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 291058 ATGGCATGATTAGGACAGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 290999

Qy    121 GTTGTCTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290998 GTTGTCTCTTAGAAATAAGGGGATAAAAATATTACCATTGATGCTAATAACATAAAAGAA 290939

Qy    181 GATCATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAATAGTACGTGCA    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290938 GATCATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAATAGTACGTGCA 290879

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTGTTGTAAGTTATATGATTTTTTAGGACCAGATTATTTGATCAA    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290878 CATCAATTGAAATTAGATGTTGTAAGTGTTGTAAGTTATATGATTTTTTAGGACCAGATTATTTGATCAA 290819

Qy    301 TATACTTCAGTAGCTGTAACTGTGTCACATGGTAAAACTTCTACAACAGGTTTATTATCA    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290818 TATACTTCAGTAGCTGTAACTGTGTCACATGGTAAAACTTCTACAACAGGTTTATTATCA 290759

Qy    361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290758 CATGTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 290699

Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTAAGT    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290698 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAAATATAGACGTCACTTTTAAGT 290639

Qy    481 TATAAACCTGATTACGCAATTAATGACAAATATTGATTCGATCATCCTGATTATTTTAAA    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290638 TATAAACCTGATTACGCAATTAATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 290579

Qy    541 GATATTAATGATGTTTTGATGCATTCCAAGAAAATGCACATAATGTAAAAAAGGTATT    600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290578 GATATTAATGATGTTTTGATGCATTCCAAGAAAATGCACATAATGTAAAAAAGGTATT 290519

Qy    601 ATTGCTTGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTATTAT    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290518 ATTGCTTGGGTGATGATGAACATCACGTAAAATTGAAGCAGATGTTCCAATTATTATTAT 290459

Qy    661 TATGGATTAAAGATTCGGATGACATTATGTCTCAAAATATTCAAATTACGGATAAAGGT    720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290458 TATGGATTAAAGATTCGGATGACATTATGTCTCAAAATATTCAAATTACGGATAAAGGT 290399

Qy    721 ACTGCTTTGATGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290398 ACTGCTTTGATGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT 290339

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA    840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290338 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 290279

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAAT    900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290278 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTAAACGTCGTTTCAAT 290219

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAGAAATT    960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290218 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACCATCCAAGAGAAATT 290159

Qy    961 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290158 AGTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 290099

Qy   1021 CAACCACACTTTCTCTAGAACACAGGCATTTTTAAATGAATTTTAAATGAATTTGCAGAAAGTTTAAGT   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290098 CAACCACACTTTCTCTAGAACACAGGCATTTTTAAATGAATTTTAAATGAATTTGCAGAAAGTTTAAGT 290039

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGCATAGATTAGAGAAAAATACTCGGCGCA   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290038 AAAGCAGATCGTGTATTCTTATGTGCATAGATTAGAGAAAAATACTCGGCGCA 289979

Qy   1141 TTAACGATACAAGATTAATTGAATAAAATGAAGGTGCATCGTAATTAATGAAGATTCT   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 289978 TTAACGATACAAGATTAATTGAATAAAATGAAGGTGCATCGTAATTAATGAAGATTCT 289919

Qy   1201 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTATTATTGGGTGCAGGTGATATT   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 289918 ATTAATGTATTAGAACAATTTGATAATGCTGTTATTTATTATTGGGTGCAGGTGATATT 289859

Qy   1261 CAAAAAATTACAAAATGCATATATTTAGATAAATTGATAAATGCATGAAAAAATGCGTTTTAA   1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 289858 CAAAAAATTACAAAATGCATATATTTAGATAAATTGATAAATGCATGAAAAAATGCGTTTTAA 289805
```

```
RESULT 6
AP003363/c
LOCUS       AP003363               342600 bp    DNA       linear   BCT 24-APR-2003
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
            sequence, section 6/9.
ACCESSION   AP003363 BA000017
VERSION     AP003363.2  GI:14247399
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
```

```
AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
          Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
          Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
          Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
          Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
          Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
          Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
          aureus
JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
MEDLINE   21311952
PUBMED    11418146
REFERENCE 2  (bases 1 to 342600)
AUTHORS   Ohta,T.
TITLE     Direct Submission
JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
          of Medical Technology and Nursing, Department of Medical
          Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
          (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
          Fax:81-298-53-3454)
COMMENT   On May 29, 2001 this sequence version replaced gi:13875943.
FEATURES             Location/Qualifiers
     source          1..342600
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /mol_type="genomic DNA"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene            complement(42..815)
                     /gene="SAV1628"
     CDS             complement(42..815)
                     /gene="SAV1628"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB57790.1"
                     /db_xref="GI:14247400"
                     /translation="MKGHQFSRNELAIGQQGLDLILKQQTVVILGVGVGVGSFAABALART
                     NIGHILILDKDDVDITNVNRQLHALTTVGQSKVSLMBERIKLINPDCKVTPLHMFYT
                     EETYEIFDNYDIDYFVDASDTIMYKVHLMKECLDRGIKVISSMGAANKTDPTRFTIS
                     DISKTHTDPMAKIIRQKLKKLGIRKGIPVVFSDESPIVIREDVKDIVGDKNAINRKGQ
                     MPPSSNAFVPSVVGLICASYVNDVLKDIPVRRIKDKGQ"
     gene            complement(1029..1130)
                     /gene="SAV1629"
     CDS             complement(1029..1130)
                     /gene="SAV1629"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57791.1"
                     /db_xref="GI:14247401"
                     /translation="MGFLAAYTPYMKDVQNKKQGTHLYITGRMIKLL"
     gene            complement(1276..3042)
                     /gene="aspS"
     CDS             complement(1276..3042)
                     /gene="aspS"
                     /note="SAV1630"
                     /codon_start=1
                     /transl_table=11
                     /product="aspartyl-tRNA synthetase"
                     /protein_id="BAB57792.1"
                     /db_xref="GI:14247402"
                     /translation="MSKRTTYCGLVTEAFLGQEITLKGWNNRRDLGGLIFVDLRDRE
                     GIVQVVFNPAFSEEALKIAETVRSEYVVEVQGTVTKRDPETVNPKIKTGQVQVTNI
                     KVINKSETPPFSINEENVNVDENIRLKRYLDLRRQELACTFKMRHQITRSIRQYLDD
                     EGFDIETPVLTKSTPEGARDYLVPSRVHDGEFYALPQSPQLFKQLLMISGFDKYQI
                     VKCFRDEDIRADRQPEFTQVDIEMSFVDQSDVMQMGEEMLKKVVKEVKGVEINGAPPR
                     MTYEKAMRRYGSDKPDTRPEMELIDVSQLGRDMDFKVFKDTVENDGEIKAIVAKGAAE
                     QYTRKDMDALTEFVNIYGAKGLAWVKVEDGLTGPIGRFPETENVETLLITGAEAGD
                     LVMFVADKPNVVAQSLGALRVLAKELGLIDETKLNFLWVTDWPLLEYDBDAKRYVAA
                     HHPFTSPKEADIAKLGTAPEEAEANAYDIVLNGYELGGGSIRIHDGELQEKMFEVLGF
                     TKEQAQEQFGFLLDAFKYGAPPHGGIALGLDRLVMLLITNRTNLRDTIAFFKTASATCL
                     LTNAPGEVSDKQLEELSLRIH"
```

```
     gene            complement(3058..4320)
                     /gene="hisS"
     CDS             complement(3058..4320)
                     /gene="hisS"
                     /note="SAV1631"
                     /codon_start=1
                     /transl_table=11
                     /product="histidyl-tRNA synthetase"
                     /protein_id="BAB57793.1"
                     /db_xref="GI:14247403"
                     /translation="MIKIPRGTQDILPEDSKKWRYIENQLDELMTFYNYKEIRTPIPE
                     STDLFARGVGDSTDVVQKEMYTFKDKGDRSITLRPGTAAVVRSYIEHKMQGNPNQPI
                     KLIYNGPMFRYERKQKGRYRQPNQFGVEAIGAENPSVDAEVLAMVMHIYQSFGLKHIK
                     LVINSVGDMASRKEYNEALVKHFPVIHEFCSDCQSRLHTNPMRILDCKVDRDKEAIK
                     TAPRITDFLNEESKAYYEQVKAYLDDLGIPYIEDNLVRGLDYYTHTAFELMMDNFNY
                     DGAITTLCGGRYNGLLELLDGPSTGIGFALSIERLLLALEEEGELDIRENLDLFI
                     VTMGDQADRYAVKLLNHLRHNGIKADKDYLQRKIKGQMKQADRLGAKFTIVIGDQELE
                     NNKIDVKNMTTGRSETIELDALVEYFKK"
     gene            complement(4782..5657)
                     /gene="lytH"
     CDS             complement(4782..5657)
                     /gene="lytH"
                     /note="SAV1632"
                     /codon_start=1
                     /transl_table=11
                     /product="N-acetylmuramoyl-L-alanine amidase"
                     /protein_id="BAB57794.1"
                     /db_xref="GI:14247404"
                     /translation="MKKIEAWLSKKGLGKRTLIVVIAFVLFIIFLFLLLNSNEDSG
                     NTITTENAELRTGPNAAYPVIYKVEKGDHFKKIGKVGKWIEVEDTSSNEKGWIAGWHT
                     NLDIVADNTKEKNPLQGKTIVLDPGHGGSDQGASSNTKYKSLEKDYTLKTAKELQRTL
                     EKEGAATVKMRTDDTTVSLENRDIKGDAYLSIHNDALESSNANGMTVYWHDNQRALA
                     DTLDATIQKKGLLSNRGSRQENYQVLRQTKVPAVLLELGYISNPTDETMIKDQLHRQI
                     LEQAIVDGLKIYFSA"
     gene            complement(5654..6106)
                     /gene="SAV1633"
     CDS             complement(5654..6106)
                     /gene="SAV1633"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB57795.1"
                     /db_xref="GI:14247405"
                     /translation="MKVVVQRVKEASVTNDTLNNQIKKGYCLLVGIGQNSTEQDADVI
                     AKKIANARLFEDDNNKLNFNIQQMNGEILSVSQFTLYADVKKGNRPGFSNSKNPDQAV
                     KIYEYFNDALRAYGLIVKTGEFGTHMNVSINNDGPVTIIYESQDGKIQ"
     gene            complement(6118..8307)
                     /gene="relA"
     CDS             complement(6118..8307)
                     /gene="relA"
                     /note="SAV1634"
                     /codon_start=1
                     /transl_table=11
                     /product="GTP pyrophosphokinase"
                     /protein_id="BAB57796.1"
                     /db_xref="GI:14247406"
                     /translation="MNNEYPYSADEVLHKAKSYLSADEYEYVLKSYHIAYEAHKGQFR
                     KNGLPYIMHPIQVAGILTEMRLDGPTIVAGFLHDVIEDTPYTFEDVKEMNEEVARIV
                     DGVTKLKKVKYRSKEEQQAENHRKLFIAIAKDVRVLVKLADRLHNMRTLKAMPREKQ
                     IRISRETLSIYAPLAHHLGINTIKWELEDTALRYIDNVQYFRIVALMKGKRSEREAYI
                     ETAIDRIRTEMDRMNIEGDINGRFKDYIYSIYRKMMKQKKQFDQIFDLLAIRVIVNSIN
                     DCYAILGLVHTLWKPMPGRFKDYIAMPKQNLYQSLHTTVGPNGDPLEIQIRTPDMHE
                     IASHGVAAHWAYKBGKKVSEKDQTYQNKLNWLKGLAEADHTSSDAQEFMETLKYDLQS
                     DKVYAFTPASDVIELPYGAVPIDFAYAIHSEVGNKMIGAKVNGKIVPIDYILQTGDIV
                     EIRTSKHSYGPSRDWLKIVKSSSAGKIKSFFKKDRSSNIEKGRMMVEVEIKEGGPR
                     VEDILTEKNIQVVNEKYNFANEDDLFAAVGFGVTSLQIVNKLTERQRILDKQRALNE
                     EQEVTKSLPIKDNIITDSGVYVEGLENVLIKLSKCCNPIPGDDIVGYITRKGHGIKVHR
                     TDCPNIKNETERLLNVEWVSKDATQKYQVLEVTAYDRNGLLAYDRYVAA
                     KVSGRSDIDKNAIINISVMVKNVNDVRVVEKIKQLGDVTVTRVWN"
     gene            complement(8735..9253)
                     /gene="apt"
     CDS             complement(8735..9253)
                     /gene="apt"
```

```
Qy    361  CATGTTATGGAATGGTGATAAAAGACTTCATTTTTTAATTGGTGATGGCACAGGTATGGGA  420
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132439  CATGTTATGGAATGGTGATAAAAAGACTTCATTTTTTAATTGGTGATGGCACAGGTATGGGA 132380

Qy    421  TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT  480
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132379  TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 132320

Qy    481  TATAAACCTGACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132319  TATAAACCTGACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA 132260

Qy    541  GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT  600
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132259  GATATTTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 132200

Qy    601  AATGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132199  AATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT 132140

Qy    661  TATGGATTTAAAGATTCCGATGACATTTATGCTCAAAATATTCAAATTACCGATAAAGGT  720
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132139  TATGGATTTAAAGATTCCGATGACATTTATGCTCAAAATATTCAAATTACCGATAAAGGT 132080

Qy    721  ACTGCTTTTGATGTATGTGGATGGTGAGTTTTATGATCCACTTCCTGTCTCCACAATAT  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132079  ACTGCTTTTGATGTATGTGGATGGTGAGTTTTATGATCCACTTCCTGTCTCCACAATAT 132020

Qy    781  GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCCATTAGTTATTGAGAGAAGCTA  840
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132019  GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCCATTAGTTATTGAGAGAAGCTA 131960

Qy    841  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT  900
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131959  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACCTCGTTTCAAT 131900

Qy    901  GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT  960
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131899  GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 131840

Qy    961  AGTGCTACAAATGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131839  AGTGCTACAAATGAAACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT 131780

Qy   1021  CAACCACCACACTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 1080
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131779  CAACCACCACACTTCTCTAGAACACAGCAAGCATTTTAAATGAATTTGCAGAAAGTTTAAGT 131720

Qy   1081  AAAGCAGATGTGTATTCTTATGTGAAATTTTTGGATCAATTGAAATTAGAGAAAATACTGGCGCA 1140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131719  AAAGCAGATGTGTATTCTTATGTGAAATTTTTGGATCAATTGAAATTAGAGAAAATACTGGCGCA 131660

Qy   1141  TTAACGATACAAGATTTAAATTGATAAAATTGAAGGTGCATCGTTAATTGAAGATTCT 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131659  TTAACGATACAAGATTTAAATTGATAAAATTGAAGGTGCATCGTTAATTGAAGATTCT 131600

Qy   1201  ATTAATGTATTAGAACAATTGATAATGTCTGTTATTTATTTATTGGGTGCAGTGGATATT 1260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131599  ATTAATGTATTAGAACAATTGATAATGTCTGTTGTTTATTTATTGGGTGCAGTGGATATT 131540

Qy   1261  CAAAAATTACAAAAATGCATATTTAGATAAATTTAGATAAATTAGGCATGAAAAATGCGTTTAA 1314
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 131539  CAAAAATTACAAAAATGCATATTTAGATAAATTTAGATAAATTAGGCATGAAAAATGCGTTTAA 131486
```

```
gene            /note="SAV1635"
                /codon_start=1
                /trans_table=11
                /product="adenine phosphoribosyl transferase"
                /protein_id="BAB57797.1"
                /db_xref="GI:14247407"
                /translation="MDLKQYVSEVQDWPKPGVSFKDITTIMDNGEAYGYATDKIVEYA
                KDRDVDIVVGPEARGFIIGCPVAYSMGIGFAPVRKEGKLPREVIRYEYDLEYGTNVLT
                MHKDAIKPGQRVLITDDLLATGGTIEAAIKLVEKLGGIVVGIAFIIELKYLNGIEKIK
                DYDVMSLISYDE"
gene            complement(9275..11548)
                /gene="SAV1636"
CDS             complement(9275..11548)
                /gene="SAV1636"
                /note="similar to single-strand DNA-specific exonuclease"
                /codon_start=1
                /trans_table=11
                /product="hypothetical protein"
                /protein_id="BAB57798.1"
                /db_xref="GI:14247408"
                /translation="MIKPKYKWKLTKPAEYISDELTSKLKLTPIVKKILESKSIIDEQ
                AIESIISDTDVNHDALQLSDMTKAIERIKRATANDEKILVYGDYDADGVTSTTILVTT
                LQLLGAQVGWHIPNRFTEGYGPNELAFRNAHDEGITLIITVDNGIQGHNEIKMVQDLG
                VDVIVTDHHEIGSTLPEAYAIVHPMHPSFNYPFQQLCGAGVAYKLAQALIENVPDYFK
                ALVAIGTIADILVSLTDENRSLVKQGLKVLNDQCPTSVKALLKEAGYNDNIDEETIGFI
                IGPRLNAVGRLDDASLACELLMTDEEAAFLAEQVEHFNRERKDIVATITEBAMAMA
                ETKVKKGDLFLLLAKENWHEGVLGIVASKIVETFALPTLILNIDREQNHAKGSARSID
                QVSMFEILSAHQELIAKFGGHHMAAGMTMDIENIESLAEGLNKWMKELSKTTSLDPVK
                GESNIAALFWQNGHLEPELQDEQPINILGSVQINEWNGNQSPQLIIQDIAMNEQQILD
                YRSKRKSLPFTENDENIVVLIHPKSDKVNANEYYYGBEIKQQTDKVLRDLPTSMEDL
                SNSLQQLQFSQLYIVLQHNHSIYFDGIPNMDIFKKCYKALITKQETNIQKEGMLLCQH
                LSVKPDTLKFMIKFMLKFVFLDLKFVTQEDGLIRINQQPDKRSIDSSKVYQLRQORMDVEKQL
                LYQDFSEIKNWIKSQLS"
gene            complement(11751..14030)
                /gene="secF"
CDS             complement(11751..14030)
                /gene="secF"
                /note="SAV1637"
                /codon_start=1
                /trans_table=11
                /product="protein-export membrane protein SecDF"
                /protein_id="BAB57799.1"
                /db_xref="GI:14247409"

Query Match    92.0%;  Score 1212;  DB 1;  Length 342600;
Best Local Similarity  99.8%;
Matches 1312;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy      1  ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC  60
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132799  ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 132740

Qy     61  ATGCATGATTAGGACATGAAGTTCAAGGATGTCGGATATTCGGATTACGTATTTTACGAA  120
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132739  ATGCATGATTAGGACATGAAGTTCAAGGATGTCGGATATTCGGATTACGTATTTTACAGAA 132680

Qy    121  GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA  180
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132679  GTTGCTCTTAGAAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 132620

Qy    181  GATANTGGTAGTTATACAAGGTTATCAAGTGGTTGTAATGCATTCGCGAGTAGCCATGAAGA 240
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132619  GATAATGGTAGTTATACAAGGTTATAATGATGTTTTAGGACAGAGAAATAGTACGTGCA 132560

Qy    241  CATCAATTGAAATTAGATGTTGTAAGTTATATTGATAATGATTTTTTAGGACAGATTATTGATCAA  300
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132559  CATCAATTGAAATTAGATGTTGTAAGTTATATTGATAATGATTTTTTAGGACAGATTATTGATCAA 132500

Qy    301  TATACTTCAGTAGCTGTAACTGGTGCACATGGTGCAAACTTCACAACAGGTTTATTATCA  360
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132499  TATACTTCAGTAGCTGTAACTGGTGCACATGGTGCAAACTTCACAACAGGTTTATTATCA 132440
```

```
                                                                                                                                Qy    301  TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACACAGGTTTATTATCA  360
                                                                                                                                          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                                                                                Db   3485  TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACACACAGGTTTATTATCA  3426

 /gene="MW1687"                                                                                                                 Qy    361  CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCCACAGGTATGGGA  420
 /note="ORFID:MW1687"                                                                                                                      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 thioredoxin homolog"                                                                                                           Db   3425  CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCCACAGGTATGGGA  3366
 /codon_start=1
 /transl_table=11                                                                                                               Qy    421  TTGCTGAAAGTGATTATTTCGCTTTGAGCATGTGAATATAGACGTCACTTTTTAAGT  480
 /protein_id="BAB95552.1"                                                                                                                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 /db_xref="GI:21204857"                                                                                                         Db   3365  TTGCTGAAAGTGATTATTTCGCTTTGAGCATGTGAATATAGACGTCACTTTTTAAGT  3306
 /translation="MKQLESEQQFESLKQGATVFEFTAGWCPDCRVIEPDLPELEARY
 PMFDFVSVDRDKFMDICIENGIMGIPSFLVYKNGELLGSYIGKERKSIEQIDAFLAQY    Qy    481  TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  540
 V"                                                                                                                                        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                                                                                Db   3305  TATAAACCTGATTACGCAATTATGACAAATATTGATTTCGATCATCCTGATTATTTTAAA  3246
 gene      complement(9612. .10688)
           /gene="MW1688"                                                                                                       Qy    541  GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTAAAAAAGTATT  600
 CDS       complement(9612. .10688)                                                                                                        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /gene="MW1688"                                                                                                       Db   3245  GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTAAAAAAGTATT  3186
           /note="ORFID:MW1688
           endo-1,4-beta-glucanase homolog"                                                                                     Qy    601  ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAT  660
           /codon_start=1                                                                                                                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /transl_table=11                                                                                                     Db   3185  ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTTATTAC  3126
           /protein_id="BAB95553.1"
           /db_xref="GI:21204858"                                                                                               Qy    661  TATGGATTAAAGATTCGGATGACATTTATCGTCAAAATATTCAAAATTACGGATAAAGGT  720
           /translation="MNINKKVTLQRIQTLTELHGAPGFEBEVKNYMTQOMAPYDEFI                                                                           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 ENRMGGFFGVKKSKNPNAKRVMIAAHMDEIGFMITNIKNGMIQFTNLGGVANDIWQG                                                                        Db   3125  TATGGATTAAAGATTCGGATGACATTTATCGTCAAAATATTCAAAATTACGGATAAAGGT  3065
 QRLVIKNRNGDKIIGVVSNIPKHFRTGSEGAPEIKDLTLDIGAQNEDEVRERGIDIGD
 TIVPHTPTQLSEHRYSAKAWDNRYGCVLAIELIELDLEDVDLYVGANVQEEVGL                                                                           Qy    721  ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  780
 RGAKASAEMIDDDVAFVVDCSPANDVKGSQPLSGELGKGTLIRIKDGTMILKPVFRDY                                                                                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 LLKLVEAHDIEHQYYMSPGGTDGGEIHKANIGIPTAVIGVCARYIHSTDSVFDIRDYF                                                                       Db   3065  ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATAT  3006
 AARSLLSEAICNLDNNQIETLQYK"
                                                                                                                                Qy    781  GGTGACCATACAGTTTAAATGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA  840
 gene      10775. .11086                                                                                                                   |||||||||||||||||||||||||||||||||||||||||||||||||||||
           /gene="MW1689"                                                                                                       Db   3005  GGTGACCATACAGTTTAAATGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA  2946
 CDS       10775. .11086
           /gene="MW1689"                                                                                                       Qy    841  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  900
           /note="ORFID:MW1689"                                                                                                           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /codon_start=1                                                                                                       Db   2945  GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGGTGTTAAACGTCGTTTCAAT  2886
           /transl_table=11
           /product="hypothetical protein"                                                                                      Qy    901  GAAACTACAATTGCAAATCAAGTTATTGTTAGATGATTATGCACACCATCCAAGAGAAATT  960
           /protein_id="BAB95554.1"                                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /db_xref="GI:21204859"                                                                                               Db   2885  GAAACTACAATTGCAAATCAAGTTATTGTTAGATGATTATGCACACCATCCAAGAGAAATT  2826
           /translation="MTKLKYIIPTIIAVAIVIISTISIIQYINRKRYNPVKVLNEVKS
 YFMNVKGSYIVEPFVHPETDKYRLVIYQGGITTIKNGQNIHYDFYADAYTGEVINIVE                                                                      Qy    961  AGTTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTTATTT  1020
 R"                                                                                                                                        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                                                                                Db   2825  AGTTGCTACAATTGAAACAGCACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTTATTT  2766
 .        gene      complement(11211. .12053)
           /gene="MW1690"                                                                                                       Qy   1021  CAACCACACACTTTCTCTAGAACACAGGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  1080
 CDS       complement(11211. .12053)                                                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /gene="MW1690"                                                                                                       Db   2765  CAACCACACACTTTCTCTAGAACACAAGCATTTTAAATGAATTTGCAGAAAGTTTAAGT  2706
           /note="ORFID:MW1690"
           /codon_start=1                                                                                                       Qy   1081  AAAGCAGATCGTGTATTCTTATGTGAAAATTTTGATCAATTAGAGAAAATACTGCGCA  1140
           /transl_table=11                                                                                                                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           /product="conserved hypothetical protein"                                                                            Db   2705  AAAGCAGATCGTGTATTCTTATGTGAAAATTTTGATCAATTAGAGAAAATACTGCGCA  2646

 Query Match       88.1%;   Score 1161;   DB 1;   Length 272850;                                                                Qy   1141  TTAACGATACAAGATTTAATTGATTAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  1200
 Best Local Similarity 99.8%;   Pred. No. 0;                                                                                               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Matches 1311; Conservative   0;  Mismatches   3;  Indels   0;  Gaps  0;                                                        Db   2645  TTAACGATACAAGATTTAATTGATTAATTGAAGGTGCATCGTTAATTAATGAAGATTCT  2586

 Qy     1  ATGACACACTATCATTTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC  60                                                        Qy   1201  ATTAATGTATTAGAACAATTGATAATGTCTGTTATTTATTTATGGTGCAGGTGATATT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                                                                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3785  ATGACACACTATCATTTTGTCGGAATTAAAGGTTCGGCATGAGTTCATTAGCACAAATC  3726                                                      Db   2585  ATTAATGTATTAGAACAATTGATAATGTCTGTTATTTATTTATGGTGCAGGTGATATT  2526

 Qy    61  ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTCGAGAACTACGTATTACAGAA  120                                                       Qy   1261  CAAAAATTACAAAATGCATATTTAGATAATAAATAGTACGTGCA  1314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                                                                     ||||||||||||||||||||||||||||||||||||||||||||
 Db   3725  ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTCGAGAACTACGTATTACAGAA  3666                                                      Db   2525  CAAAAATTACAAAATGCATATTTAGATAATAAATAGGCATGAAAAATGCGTTTTAA  2472

 Qy   121  GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTGATGCTAATAACATAAAAGAA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3665  GTTGCTCTTAGAAATAAGGGGATAAAATATTACCATTGATGCTAATAACATAAAAGAA  3606                                                     RESULT 8
                                                                                                                               AF034076
 Qy   181  GATATGGTAGTTATACAAGGTAATGCATTCCGCGAGTAGCCATGAAGAAAAGTACGTGCA  240                                                     LOCUS        AF034076      1314 bp    DNA      linear    BCT 16-JUL-1999
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                                                         DEFINITION   Staphylococcus aureus UDP-N-acetylmuramoyl-L-alanine synthetase
 Db   3605  GATATGGTAGTTATACAAGGTAATGCATTCCGCGAGTAGCCATGAAGAAAAGTACGTGCA  3546

 Qy   241  CATCCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3545  CATCCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTTAGGACAGATTATTGATCAA  3486
```

```
ACCESSION   AF034076          (murC) gene, complete cds.
VERSION     AF034076.1  GI:2642658
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 1314)
  AUTHORS   Lowe,A.M. and Deresiewicz,R.L.
  TITLE     Cloning and sequencing of Staphylococcus aureus murC, a gene
            essential for cell wall biosynthesis
  JOURNAL   DNA Seq. 10 (1), 19-23 (1999)
  MEDLINE   20029270
  PUBMED    10565540
REFERENCE   2  (bases 1 to 1314)
  AUTHORS   Lowe,A.M. and Deresiewicz,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1997) Channing Laboratory, Brigham and Women's
            Hospital and Harvard Medical School, 181 Longwood Ave, Boston, MA
            02115, USA
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="Staphylococcus aureus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1280"
     gene            1..1314
                     /gene="murC"
     CDS             1..1314
                     /gene="murC"
                     /function="cell wall biosynthesis"
                     /note="MurC; UDP-N-acetylmuramate-alanine ligase"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylmuramoyl-L-alanine synthetase"
                     /protein_id="AAB87090.1"
                     /db_xref="GI:2642659"
                     /translation="MTHYHFVGIKGSGMSSLAQIMHDLGHEVQGSDIENYVFTEVALR
                     NKGIKILPFDANNIKEDMVVIQGNAFASSHEEKARAHQMKLDVVSYNDFLGQIIDQYT
                     SVAVTGAHGKTSTGLLSHVMNGDKKTSFLIGDGTGMGLPESDYFAFEACEYRRHFLS
                     YKPDYAIMTNIDFDRPDYFKDINDVFDAFQEMAHNVKKGIIAWGDDEHLRKIEADVPI
                     YYYGFKDSDDIYAQNIQITDKGTAPVYVDGEFYDHFLSPQYGDHTVLNALAVIAISY
                     LEKLDVTNIKEALEFPGGVKRFNETTIANQVVIDYAHHPREISATIETARKKYPHK
                     EVVAVFQPHTFSRTQAFLNEFAESLSKADRVFLCEIFGSIRENTGALTIQDLIDKIEG
                     ASLNEDSINVLEQFDNAVVLFKGAGDIQKLQNAYLDKLGMKNAF"
ORIGIN

Query Match     75.5%;  Score 995;  DB 1;  Length 1314;
Best Local Similarity   99.7%;
Matches 1195; Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
                     Pred. No. 0;
```

```
Qy     1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGACACACTATCATTTTGTCGGAATTAAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60

Qy    61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATGCATGATTTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120

Qy   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTTGCTCTTAGAAATAAGGGGATAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180

Qy   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATATGGTAGTTATACAAGGTAATGCATTCGCGAGTAGCCATGAAGAAAATAGTACGTGCA 240

Qy   241 CATCAATTGAAATTAGGACGTTGTAAGTTATATAATGATTTTTAGGACAGATTATTGATCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCAATGAAATTAGGACGTTGTAAGTTATATAATGATTTTTAGGACAGATTATTGATCAA 300

Qy   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
```

```
Qy   361 CATGTTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CATGTTTATGAATGGTGATAAAAAGACTTCATTTTTAATTGGTGATGGCACAGGTATGGGA 420

Qy   421 TTGCCTGAAAGTGATTATTTCGCTTTGAGGCATGTGATATAGACGTCACTTTTAAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTGCCTGAAAGTGATTATTTCGCTTTGAGGCATGTGATATAGACGTCACTTTTAAGT 480

Qy   481 TATAAACCTGATTACGCAATATGACAAATATGATTTCGATCATCCTGATTATTTAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TATAAACCTGATTACGCAATATGACAAATATGATTTCGATCATCCTGATTATTTAAA 540

Qy   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAAGCTATT 600

Qy   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTATTAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTATTAT 660

Qy   661 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAATATTCAAATTACGGATAAAGGT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATGGATTAAAGATTCGGATGACATTTATGTCTCAAATATTCAAATTACGGATAAAGGT 720

Qy   721 ACTGCTTTTGATGTGTATGGTGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTGCTTTTGATGTGTATGGTGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy   781 GGTGACCACATAGTTTAAATGCATTCGGATTAGTCGTTATTAGAGAAGCTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGTGACCACATAGTTTAAATGCATTCGGATTAGTCGTTATTAGAGAAGCTA 840

Qy   841 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTGGTGGTGTTAAACGTCGTTCAAT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATGTTACAAATATTAAAGAAGCATTAGAAAACGTTGGTGGTGTTAAACGTCGTTCAAT 900

Qy   901 GAAACTACAATTGCAAATCAAGTTATTGTGAATGATATTGCACCCATCCAAGAGAAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAACTACAATTGCAAATCAAGTTATTGTGAATGATATTGCACCCATCCAAGAGAAAAT 960

Qy   961 AGTGTCACAATTGAAACAGCACAGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGTGTCACAATTGAAACAGCACAGAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020

Qy  1021 CAACCACACTTCTCTAGAACACAGGCATTTTAATTGCAGAAAGTTTAAGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CAACCACACTTCTCTAGAACACAGGCATTTTAATTGCAGAAAGTTTAAGT 1080

Qy  1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140

Qy  1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTC 1199
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TTAACGATACAAGATTTAATTGATAAAATTGAAGGTGCATCGTTAATTAATGAAGATTC 1199
```

```
RESULT 9
AX191741
LOCUS       AX191741                1314 bp    DNA     linear     PAT 15-AUG-2001
DEFINITION  Sequence 23 from Patent WO0149775.
ACCESSION   AX191741
VERSION     AX191741.1  GI:15209910
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 23 12-JUL-2001;
            Avi Biopharma, Inc. (US)
```

```
FEATURES        Location/Qualifiers
    source      1..1314
                /organism="Staphylococcus aureus"
                /mol_type="unassigned DNA"
                /db_xref="taxon:1280"
ORIGIN

Query Match        75.5%;   Score 995;  DB 6;  Length 1314;
Best Local Similarity   99.7%;   Pred. No. 0;
Matches 1195;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy      1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTGTTCATTAGCACAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTGTTCATTAGCACAAATC 60

Qy     61 ATGCATGATGAGTTAGGACATGAAGTTCAAGGATCCGATATTGAGAACTAGGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGCATGATGAGTTAGGACATGAAGTTCAAGGATCCGATATTGAGAACTAGGTATTTACAGAA 120

Qy    121 GTTGCTCCTTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCTCCTTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAAGAA 180

Qy    181 GATAATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAATAGTACGTGCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATAATGGTAGTTATACAAGGTAATGCATTCGCCGAGTAGCCATGAAGAAAAAGCACGTGCA 240

Qy    241 CATCAATTGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CATCAAATGAAATTAGATGTTGTAAGTTATAATGATTTTTAGGACAGATTATTGATCAA 300

Qy    301 TATACTTCAGTAGCTGTCAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATACTTCAGTAGCTGTCAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360

Qy    361 CATGTTATGAAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGCACAGGTATGGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATGTTATGAAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGCACAGGTATGGGA 420

Qy    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTGCCTGAAAGTGATTATTTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTTAAGT 480

Qy    481 TATAAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATAAAACCTGATTACGCAATTATGACAAATATTGATTCGATCATCCTGATTATTTTAAA 540

Qy    541 GATATTAATGATGTTTTGATGCAATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GATATTAATGATGTTTTGATGCAATTCCAAGAAATGGCACATAATGTTAAAAAAGGTATT 600

Qy    601 ATTGCTTGGGGTGATGAACAATCTACGTCACGTAAATTACGGATAAAGGT 660
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATTGCTTGGGGTGATGAACAATCTACGTCACGTAAATTACGGATAAAGGT 660

Qy    661 TATGGATTTAAAGATTCGGATGACATTTATGTCTAAATATTCAAAATATTCAAATTATTAT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TATGGATTTAAAGATTCGGATGACATTTATGTCTAAATATTCAAAATATTCAAATTATTAT 720

Qy    721 ACTGCTGTTTGATGTGATGGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACTGCTGTTTGATGTGATGGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTCGATTAGTTATTTAGAGAAGCTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGTGACCATACAGTTTAAATGCATTAGCTCGATTAGTTATTTAGAGAAGCTA 840

Qy    841 GATGTTACAAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GATGTTACAAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 900

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960

Qy    961 AGTGCTACAATTGAAACAGCAGCACGACGAAAATATCCACATAAAGAGAGTTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGTGCTACAATTGAAACAGCAGCACGACGAAAATATCCACATAAAGAGAGTTGTTGCAGTATTT 1020

Qy   1021 CAACCACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAACCACCACACTTTCTCTAGAACACAGGCATTTTAAATGAATTGCAGAAAGTTTAAGT 1080

Qy   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGCAGATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAGAGAAAATACTGGCGCA 1140

Qy   1141 TTAACGATACAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTC 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTAACGATACAGATTTAATTGATAAAAATTGAAGGTGCATCGTTAATTAATGAAGATTC 1199


RESULT 10
E35613
LOCUS       E35613             1351 bp    DNA        linear   PAT 18-JUN-2001
DEFINITION  MurC.
ACCESSION   E35613
VERSION     E35613.1   GI:13019102
KEYWORDS    JP 1999225773-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1351)
  AUTHORS   Nicola,G.W. and Martin,K.R.B.
  TITLE     MurC
  JOURNAL   Patent: JP 1999225773-A/1 24-AUG-1999;
            SMITHKLINE BEECHAM CORP; SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT     OS   Unidentified
            PN   JP 1999225773-A/1
            PD   24-AUG-1999
            PF   03-JUL-1998 JP 1998225115
            PR   03-JUL-1997 US   60/052720
            PI   NICOLA G WALLIS, MARTIN K R BURNHAM
            PC   C12N15/09,A61K31/70,A61K38/53,A61K38/55,A61K39/085,A61K39/395,
            PC   A61K39/395,
            PC   A61K48/00,A61K48/00,C07K16/40,C12N1/21,C12P21/02,C12P21/08, PC
            C1201/68,
            PC   G01N33/566//(C12N15/09,C12R1:445),(C12N1/21,C12R1:445)
            C12N15/00,A61K37/60,
            PC   A61K37/64,(C12N15/00,C12R1:445), PC
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key            Location/Qualifiers
            FT   source         1..1351
            FT                  /organism='Unidentified'.
FEATURES    Location/Qualifiers
    source  1..1351
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN

Query Match        63.9%;   Score 842;  DB 6;  Length 1351;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 1042;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy      1 ATGACACCATCATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTGTTCATTAGCACAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGACACACTATCATTTGTCGGAATTAAAGGTTCTGGCATGAGTGTTCATTAGCACAAATC 81

Qy     61 ATGCATGATGATTAGGACATGAAGTTCAAGGATCGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 ATGCATGATGATTAGGACATGAAGTTCAAGGATCGAAGTTCAAGGATCGGATATTGAGAACTACGTATTTACAGAA 141

Qy    121 GTTGCTCTTAGAAATAAGGGGATAAAAATATATTACCATTTGATGCTAATAACATAAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCTCTTAGAAATAAGGGGATAAAAATATATTACCATTTGATGCTAATAACATAAAAGAA 180
```

```
REFERENCE    1  (bases 1 to 1351)
  AUTHORS    Wallis,N.G. and Burnham,M.K.R.
  TITLE      MurC
  JOURNAL    Patent: JP 2002300888-A 1 15-OCT-2002;
             SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM PLC
COMMENT      OS   Unidentified
             PN   JP 2002300888-A/1
             PD   15-OCT-2002
             PF   25-DEC-2001 JP 2001391079
             PR   03-JUL-1997 US    60/052720
             PI   NICOLA G WALLIS,MARTIN K R BURNHAM
             PC   C12N15/09,A61K31/7088,A61K38/43,A61K38/55,A61K45/00,A61K48/00,
             PC   A61P1/00,
             PC   A61P7/02,A61P9/00,A61P11/00,A61P13/02,A61P13/12,A61P17/00, PC
             A61P17/02,
             PC   A61P19/02,A61P19/08,A61P21/00,A61P25/00,A61P27/00,A61P29/00,
             PC   A61P31/04,
             PC   A61P39/02,A61P43/00,C07K14/195,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
             PC   C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00, PC
             C12N5/00,
             PC   A61K37/48,A61K37/64
             CC   MurC
FEATURES     Key             Location/Qualifiers
  FH
  FT   source               1..1351
  FT                        /organism='Unidentified'.

FEATURES             Location/Qualifiers
     source          1..1351
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match             63.9%;  Score 842;  DB 6;  Length 1351;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 1042;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy    1 ATGACACACTATCATTTGTCGGAATAAGGTTCTGGCATGAGTTCATTAGCACAAATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGACACACTATCATTTGTCGGAATAAGGTTCTGGCATGAGTTCATTAGCACAAATC 81

Qy   61 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 ATGCATGATTAGGACATGAAGTTCAAGGATCGGATATTGAGAACTACGTATTACAGAA 141

Qy  121 GTTGCTCTTAGAAATAAGGGGATAAAAAATATTACCATTTGATGCTAATAACATAAAGAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  142 GTTGCTCTTAGAAATAAGGGGATAAAAAATATTACCATTTGGTGCTAATAACATAAAGAA 201

Qy  181 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAATAGTACGTGCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  202 GATATGGTAGTTATACAAGGTAATGCATTCGGCGAGTAGCCATGAGAAATAGTACGTGCA 261

Qy  241 CATCAATTGAAATTGAAATGATGTTGTAAGTTATAGTTGTAAGTGTCATAATTGATCA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 CATCAATTGAAATTGAAATGATGTTGTAAGTTATAGTTGTAAGTGTCATAATTGATCA 321

Qy  301 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  322 TATACTTCAGTAGCTGTAACTGGTGCACATGGTAAAACTTCACAACAGGTTTATTATCA 381

Qy  361 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  382 CATGTTATGAATGGTGATAAAAAGACTTCATTTTAATTGGTGATGGCACAGGTATGGGA 441

Qy  421 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  442 TTGCCTGAAAGTGATTATTCGCTTTTGAGGCATGTGAATATAGACGTCACTTTTAAGT 501

Qy  481 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTTTAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  502 TATAAACCTGATTACGCAATTATGACAAATATTCGATCATCCTGATTATTTTCAAA 561

Qy  541 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAGGTATT 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  562 GATATTAATGATGTTTTGATGCATTCCAAGAAATGGCACATAAATGTTAAAAAGGTATT 621

Qy  601 ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTTATTAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAAATTGAAGCAGATGTTCCAATTTATTAC 681

Qy  661 TATGGATTAAAGATTCGGATGCATGATTATGCTCAAAATTACGGATAAAGGT 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  682 TATGGATTAAAGATTCGGATGCATGATTATGCTCAAAATTACGGATAAAGGT 741

Qy  721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy  781 GGTGACCATACAGTTTTAAATGCCATAGTCGTGTAAATTGCGATTAGTTATTGAGAAGCTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  802 GGTGACCATACAGTTTTAAATGCCATAGTCGTGTAAATTGCGATTAGTTATTGAGAAGCTA 861

Qy  841 GATGTTACAAATATTAAAGAAGCATTAAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  862 GATGTTACAAATATTAAAGAAGCATTAAGAAACGTTTGGTGGTGTAAACGTCGTTTCAAT 921

Qy  901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATATGCGACACCATCCAAGAGAAATT 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  922 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATATGCGACACCATCCAAGAGAAATT 981

Qy  961 AGTGCTGCAATTGAAACAGCACGAAAGAAATCCACATAAAGAAGTTGTTGCAGTATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  982 AGTGCTGCAATTGAAACAGCACGAAAGAAATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy 1021 CAACCACACACTTTCTCTAGAACACA 1046
         ||||||||||||||||||||||||||
Db 1042 CAACCACACACTTTCTCTAGAACACA 1067


RESULT 11
BD178030

LOCUS       BD178030      1351 bp    DNA      linear    PAT 16-APR-2003
DEFINITION  MurC.
ACCESSION   BD178030
VERSION     BD178030.1  GI:30015294
KEYWORDS    JP 2002300888-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
```

```
Qy    541 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GATATTAATGATGTTTTTGATGCATTCCAAGAAATGGCACATAATGTTAAAAAGGTATT 621

Qy    601 ATTGCTTGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTAT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ATTGCTTGGGGTGATGATGAACATCTACGTAAAATTGAAGCAGATGTTCCAATTATTATTAC 681

Qy    661 TATGGGATTAAAGATTCGGATGACGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TATGGGATTAAAGATTCGGATGACGACATTTATGCTCAAAATATTCAAATTACGGATAAAGGT 741

Qy    721 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ACTGCTTTTGATGTGTATGTGGATGGTGAGTTTTATGATCACTTCCTGTCTCCACAATAT 801

Qy    781 GGTGACCATACAGTTTTAAAATGCGATTAGCTGTAATTGGCGATTAGTTAGTTAGAGAAGCTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 GGTGACCATACAGTTTTAAAATGCGATTAGCTGTAATTGGCGATTAGTTAGTTAGAGAAGCTA 861

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTTGGTGGTGTTAAACGTCGTTTCAAT 921

Qy    901 GAAACTACAAGTTGCAAATCAAGTTATGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GAAACTACAAGTTGCAAATCAAGTTATGTAGATGATTATGCACACCATCCAAGAGAAATT 981

Qy    961 AGTGCTACAAATTGAAACAGCACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 AGTGCTACAAATTGAAACAGCACAGCACGAAAGAAAATATCCACATAAAGAAGTTGTTGCAGTATTT 1041

Qy   1021 CAACCACACACTTTCTCTAGAACACA 1046
          ||||||||||||||||||||||||
Db   1042 CAACCACACACTTTCTCTAGAACACA 1067
```

```
RESULT 12
E35614
LOCUS          E35614                 660 bp    DNA        linear    PAT 18-JUN-2001
DEFINITION     MurC.
ACCESSION      E35614
VERSION        E35614.1  GI:13019103
KEYWORDS       JP 1999225773-A/2.
SOURCE         unidentified
  ORGANISM     unidentified
               unclassified.
REFERENCE      1  (bases 1 to 660)
  AUTHORS      Nicola,G.W. and Martin,K.R.B.
  TITLE        MurC
  JOURNAL      Patent: JP 1999225773-A 2 24-AUG-1999;
               SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT        OS   Unidentified
               PN   JP 1999225773-A/2
               PD   24-AUG-1999
               PF   03-JUL-1998 JP 1998225115
               PR   03-JUL-1997 US  60/052720
               PI   NICOLA G WALLIS,MARTIN K R BURNHAM
               PC   C12N15/09,A61K31/70,A61K38/53,A61K38/55,A61K39/085,A61K39/395,
               PC   A61K39/395,
               PC   A61K48/00,A61K48/00,C07K16/40,C12N1/21,C12P21/02,C12P21/08, PC
               C12Q1/68,
               PC   G01N3/566//(C12N15/09,C12R1:445),(C12N1/21,C12R1:445), PC
               C12N15/00,A61K37/60,
               PC   A61K37/64,(C12N15/00,C12R1:445)
               CC   Strandedness: Double;
               CC   Topology: Linear;
FH             Key             Location/Qualifiers
FT             source          1..660
FT                             /organism='Unidentified'.
                               /organism="unidentified"
```

```
Qy    728 TTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGGTGACC 787
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TTGATGTGTATGTGGATGGTGAGTTTATGATCACTTCCTGTCTCCACAATATGGTGACC 122

Qy    788 ATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTATTAGAGAAGCTAGAGATGTTA 847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 ATACAGTTTTAAATGCATTAGCTGTAATTGCGATTAGTATTAGAGAAGCTAGAGATGTTA 182

Qy    848 CAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTTAAAACTCGTTCAATGAAACTA 907
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CAAATATTAAAGAAGCATTAGAAACGTTTGGTGTGTTAAAACTCGTTCAATGAAACTA 242

Qy    908 CAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGCTA 967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATTAGTGCTA 302

Qy    968 CAATTGAAACAGCACGACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTCAACCAC 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CAATTGAAACAGCACGACGAAAGAAATATCCACATAAAGAAGTTGTTGCAGTATTCAACCAC 362

Qy   1028 ACACTTTCTCTAGAACACAGCAGGCATTTTAAATGAATTTTGATCAATTAGAGAGAAAATACTGGCGCATTAAGA 1087
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 ACACTTTCTCTAGAACACAGCAGGCATTTTAAATGAATTTTGATCAATTAGAGAGAAAATACTGGCGCATTAAGA 422

Qy   1088 ATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAAATTAATGATGAAGATTCTATTAATG 1147
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 ATCGTGTATTCTTATGTGAAATTTTTGGATCAATTAAATTAATGATGAAGATTCTATTAATG 482

Qy   1148 TACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATCGTTAATTGATAAAATTGAAGGTGCATCGTAATCGTTAATTGATAAGATTCTATTAATG 1207
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    483 TACAAGATTTAATTGATAAAATTGAAGGTGCATCGTAATCGTTAATTGATAAAATTGAAGGTGCATCGTAATCGTTAATTGATAAGATTCTATTAATG 542

Qy   1208 TATTAGAACAATTTGATAATGCTCGTTTATTTTTATTTATGGTGCAGGTGATATTCAAAAT 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TATTAGAACAATTTGATAATGCTCGTTTATTTTTATTTATGGTGCAGGTGATATTCAAAAT 602

Qy   1268 TACAAAATGCATATTTAGATAAATTAGGCATGAAAATGCGTTTTAA 1314
          ||||||||||||||||||||||||||||||||||||||||
Db    603 TACAAAATGCATATTTAGATAAATTAGGCATGAAAATGCGTTTTAA 649
```

```
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32644"

ORIGIN

Query Match       32.9%;   Score 434;  DB 6;  Length 660;
Best Local Similarity   99.5%;   Pred. No. 5e-217;
Matches  584;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;
```

```
RESULT 13
BD178031
LOCUS          BD178031               660 bp    DNA        linear    PAT 16-APR-2003
DEFINITION     MurC
ACCESSION      BD178031
VERSION        BD178031:1  GI:3001595
KEYWORDS       JP 2002300888-A/2.
SOURCE         unidentified
  ORGANISM     unidentified
               unclassified.
REFERENCE      1  (bases 1 to 660)
  AUTHORS      Wallis,N.G. and Burnham,M/K.R.
  TITLE        MurC
  JOURNAL      Patent: JP 2002300888-A 2 15-OCT-2002;
               SMITHKLINE BEECHAM CORP, SMITHKLINE BEECHAM PLC
COMMENT        OS   Unidentified
               PN   JP 2002300888-A/2
               PD   15-OCT-2002
               PF   25-DEC-2001 JP 2001391079
               PR   03-JUL-1997 US  60/052720
               PI   NICOLA G WALLIS,MARTIN K R BURNHAM
               PC   C12N15/09,A61K31/7088,A61K38/43,A61K38/55,A61K45/00,A61K48/00,
               PC   A61P1/00,
               PC   A61P7/02,A61P9/00,A61P11/00,A61P13/00,A61P13/12,A61P17/00, PC
```

Left column

```
        A61P17/02,
    PC  A61P19/02,A61P19/08,A61P21/00,A61P25/00,A61P27/00,A61P29/00,
    PC  A61P31/04,
    PC  A61P39/02,A61P43/00,C07K14/195,C12N1/15,C12N1/19,C12N1/21, PC
        C12N5/10,
    PC  C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00, PC
        C12N5/00,
    PC  A61K37/48,A61K37/64
    CC  MurC
    FH  Key              Location/Qualifiers
    FT  source           1..660
    FT                   /organism="unidentified"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:32644"
ORIGIN

Query Match             32.9%;  Score 434;  DB 6;  Length 660;
Best Local Similarity   99.5%;  Pred. No. 5e-217;
Matches  584;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    728 TTGATGTGTATGTGAGTGATGTGAGTTTTATGATCATCACTTCCGTCTCCACAATATGTGACC 787
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     63 TTGATGTGTATGTGAGTGATGTGAGTTTTATGATCACTTCCGTCTCCACAATATGTGACC 122

Qy    788 ATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTAGATGTTA 847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 ATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTAGATGTTA 182

Qy    848 CAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTCAATGAAACTA 907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTCAATGAAACTA 242

Qy    908 CAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCACCATCCAAGAGAAATTAGTGCTA 967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCACCATCCAAGAGAAATTAGTGCTA 302

Qy    968 CAATTGAAACAGCAGCAGGAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTTCAACAC 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CAATTGAAACAGCAGCAGGAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTTCAACAC 362

Qy   1028 ACACTTTCTCTAGAACACACAGGCATTTTTAAATGAATTTTAAATGAATTTCAGAAGAAGTTTAAGTGAAAGCAG 1087
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 ACACTTTCTCTAGAACACACAGGCATTTTTAAATGAATTTTAAATGAATTTCAGAAGAAGTTTAAGTAAAGCAG 422

Qy   1088 ATCGTGTGATTCTTATGTGAAATTTTTCGGATCAATTAGAGAAAATACTCGGCGCCATTAACGA 1147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 ATCGTGTGATTCTTATGTGAAATTTTTCGGATCAATTAGAGAAAATACTCGGCGCCATTAACGA 482

Qy   1148 TACAAGATTTAATTGATAAAATTGAAGTGCATCGTCGTTAATTAATGAAGATTCTATTAATG 1207
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 TACAAGATTTAATTGATAAAATTGAAGTGCATCGTCGTTAATTAATGAAGATTCTATTAATG 542

Qy   1208 TATTAGAACAATTTGATAATGCTGTATTTATTATTTGGGTGCAGTGCAGTGATATTCAAAAAT 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TATTAGAACAATTTGATAATGCTGTGTGTTTTATTATTTGGGTGCAGTGCAGTGATATTCAAAAAT 602

Qy   1268 TACAAAATGCATATTTAGATGATAAAATTAGGCATGAAAAATGCGTTTTAA 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 TACAAAATGCATATTTAGATGATAAAATTAGGCATGAAAAATGCGTTTTAA 649

RESULT 14
AR194630/c
LOCUS       AR194630      619 bp   DNA   linear   PAT 20-APR-2002
DEFINITION  Sequence 179 from patent US 6348582.
ACCESSION   AR194630
VERSION     AR194630.1  GI:20241222
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unknown.
```

Right column

```
            Unclassified.
            1 (bases 1 to 619)
REFERENCE
AUTHORS     Black.M.Terence., Hodgson,J.Edward., Knowles,D.Justin.Charles.,
            Reichard,R.Winfield., Nicholas,R.O., Burnham,M.Karl.Russel.,
            Pratt,J.M., Rosenberg,M., Ward,J.M. and Lonetto,M.Arthur.
TITLE       Prokaryotic polynucleotides polypeptides and their uses
JOURNAL     Patent: US 6348582-A 179 19-FEB-2002;
FEATURES    Location/Qualifiers
source      1..619
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match             20.9%;  Score 275;  DB 6;  Length 619;
Best Local Similarity   99.7%;  Pred. No. 3.5e-133;
Matches  325;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    721 ACTTGCTTTTGATGTGTATGTGGGATGGTGAGTTTATGATCACTTCCGTCTCCACAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    611 ACTTGCTTTTGATGTGTATGTGGGATGGTGAGTTTATGATCACTTCCGTCTCCACAATAT 552

Qy    781 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 GGTGACCATACAGTTTAAATGCATTAGCTGTAATTGCGATTAGTTATTTAGAGAAGCTA 492

Qy    841 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 GATGTTACAAATATTAAAGAAGCATTAGAAACGTTGGTGTGTTAAACGTCGTTTCAAT 432

Qy    901 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 GAAACTACAATTGCAAATCAAGTTATTGTAGATGATTATGCACACCATCCAAGAGAAATT 372

Qy    961 AGTGCTACAATTGAAACAGCAGCAGGAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 AGTGCTACAATTGAAACAGCAGCAGGAAGAAATATCCACATAAAGAGAGTGTTGCAGTATTT 312

Qy   1021 CAACCACACACTTTCTCTAGAACACA 1046
          ||||||||||||||||||||||||||
Db    311 CAACCACACACTTTCTCTAGAACACA 286

RESULT 15
AR149358/c
LOCUS       AR149358      888 bp   DNA   linear   PAT 08-AUG-2001
DEFINITION  Sequence 61 from patent US 6228588.
ACCESSION   AR149358
VERSION     AR149358.1  GI:15113949
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
            1 (bases 1 to 888)
REFERENCE
AUTHORS     Benton,B., Lee,V.J., Malouin,F., Martin,P.K., Schmid,M.B. and
            Sun,D.
TITLE       Methods of screening for compounds active on Staphylococcus aureus
            target genes
JOURNAL     Patent: US 6228588-A 61 08-MAY-2001;
FEATURES    Location/Qualifiers
source      1..888
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match             6.8%;  Score 90;  DB 6;  Length 888;
Best Local Similarity   100.0%;  Pred. No. 1.2e-35;
Matches  90;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1149 ACAAGATTTAATTGATAAAATTGAAGTGCATCGTTAATTAATGAAGATTCTATTAATGT 1208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 ACAAGATTTAATTGATAAAATTGAAGTGCATCGTTAATTAATGAAGATTCTATTAATGT 508

Qy   1209 ATTAGAACAATTGATAAATGCTGTTATTTT 1238
```

us-10-712-713-1.oligo.rge

Mon Jun 7 10:15:18 2004

Db   507 ATTAGAACAATTTGATAATGCTGTTATTT 478

Search completed: June 6, 2004, 06:04:07
Job time : 5363 secs